THE UNIVERSITY OF ALBERTA

SELF-ASSESSED HEALTH OF THE ELDERLY: VALIDATION OF A BEHAVIOURALLY BASED MEASURE OF PERCEIVED HEALTH STATUS

by

ALLISON LOIS MCKINNON

(C)

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES AND RESEARCH

IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE

OF MASTER OF EDUCATION

DEPARTMENT OF EDUCATIONAL PSYCHOLOGY

EDMONTON, ALBERTA

FALL 1983

UNIVERSITY OF WASHINGTON
SEATTLE, WASHINGTON 98195
SCHOOL OF PUBLIC HEALTH AND COMMUNITY MEDICINE
Department of Health Services SC-37
Community Medicine (206) 543-7952

May 26, 1982

Ms. Allison L. McKinnon
10805 - 73 Avenue
Edmonton, Alberta
CANADA
T6E 1C8

Dear Ms. McKinnon:

The research that you propose in your letter of May 5 seems completely appropriate for the Sickness Impact Profile. Please consider this letter as an indication of permission to use the SIP in your Master's thesis research. I would, however, appreciate receiving a detailed description of the proposed research and, when it is completed, a copy of the thesis itself.

Sincerely yours,

Marilyn Bergner, Ph.D.
Associate Professor

MB:lah
Enclosures

THE UNIVERSITY OF ALBERTA

RELEASE FORM

NAME OF AUTHOR ALLISON LOIS MCKINNON

TITLE OF THESIS SELF-ASSESSED HEALTH OF THE ELDERLY: VALIDATION OF A BEHAVIOURALLY BASED MEASURE OF PERCEIVED HEALTH STATUS

DEGREE FOR WHICH THESIS WAS PRESENTED MASTER OF EDUCATION

YEAR THIS DEGREE GRANTED FALL 1983

Permission is hereby granted to THE UNIVERSITY OF ALBERTA LIBRARY to reproduce single copies of this thesis and to lend or sell such copies for private, scholarly or scientific research purposes only.

The author reserves other publication rights, and neither the thesis nor extensive extracts from it may be printed or otherwise reproduced without the author's written permission.

(SIGNED) Allison L. McKinnon

PERMANENT ADDRESS:

P.O. Box 284

Lundar, Manitoba

R0C 1Y0

DATED July 12 1983

THE UNIVERSITY OF ALBERTA

FACULTY OF GRADUATE STUDIES AND RESEARCH

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research, for acceptance, a thesis entitled SELF-ASSESSED HEALTH OF THE ELDERLY: VALIDATION OF A BEHAVIOURALLY BASED MEASURE OF PERCEIVED HEALTH STATUS submitted by ALLISON LOIS MCKINNON in partial fulfilment of the requirements for the degree of MASTER OF EDUCATION.

.....

Supervisor

.....

.....

Date....July 12, 1983.....

Dedication

Dedicated in memory of my grandmothers, Alice Bell McKenzie McKinnon and Christina Louisa Carroll Wray.

Abstract

In recognition of the need for comprehensive, valid, and reliable measures of perceived health status, expressed in terms of function, and suitable for use with the elderly, a validation study of the Sickness Impact Profile (SIP) as developed by Bergner et al. (1981) was conducted with a sample whose ages ranged from 65 to 93 years.

Thirty people attending a geriatric Day Hospital located in the city of Edmonton volunteered for the study. There were 13 males and 17 females. The predominance of chronic disease states and a tendency to multiple health problems was evident by the nature and number of diagnoses made by the physician of the Day Hospital for members of the sample. Participants were assessed on two occasions using measures of perceived health status as indicated by responses to the SIP and self-ratings of health, and of objective health status as indicated by the active problems listed in each respondent's medical record. Following the second occasion, participants in the study and health care professionals of the Day Hospital completed questionnaires in which they gave their opinions about the SIP.

In general, the SIP was able to identify and describe in detail the physical and psychosocial functional status of the sample. Test-retest reliability (r=0.80) and internal consistency (r=0.93) were high. Most members of the sample considered the SIP to give an accurate picture of how they saw their health, and and evaluation of the SIP by the

health care professionals indicated a high degree of acceptability in this clinical setting.

As a measure of health status, and specifically of perceived health, the construct validity of the SIP was supported by its significant relationships with the number of active problems listed and self-ratings of health. The SIP appeared sensitive not only to the relative severity of dysfunction of the respondents, but also to the nature of health problems experienced. In particular, when participants were divided into groups on the basis of presence or absence of diagnosed depression, the SIP discriminated between them specifically in terms of category scores in the psychosocial dimension. Notably, the age and sex of respondents were not significantly related to health status indicators in this study.

It was concluded that utilization of behaviourally based measures of perceived health such as the SIP may clarify perceived needs for services, serve as an outcome measure of treatment programs, and encourage further steps toward understanding the interrelationships between health, behaviour, and aging.

Acknowledgments

I would like to extend my sincere appreciation to Dr. T. O. Maguire for his teaching, guidance, and patience throughout the course of my program. I would also like to thank Dr. S. M. Abu-Laban and Dr. L. Wilgosh for their willingness to serve on my thesis committee.

A debt of gratitude is owed to Dr. R. Skelton, Dr. M. Watson, Mrs. D. Hawreschuk, and the staff of the Youville Day Hospital for their enthusiastic support of this research project, and I remain especially grateful to the patients of the Day Hospital for their participation in the study.

I also wish to thank R. Scott Stacey for his expert assistance with computer programming, and Joanne Pagnucco for preparation of the Textform tables. To my family and friends who provided encouragement throughout my studies, I express my deepest appreciation.

Table of Contents

Chapter		Page
I.	Introduction	1
II.	Background to the Study	3
	A. Selected Studies of Self-Rated Health	5
	B. The Sex Morbidity Ratio	10
	Interpretations	12
	C. The Performance Conception of Health	17
III.	The Sickness Impact Profile	23
	Purpose	23
	Development, Scaling, and Scoring of the SIP	24
	Reliability	25
	Validity	26
IV.	Methods	29
	A. The Sample	29
	B. Research Design	30
	C. Instruments	31
	Physician Diagnoses	31
	Active Problems Listed in Problem Oriented Medical Records	32
	Self-Ratings of Health	32
	The Sickness Impact Profile	33
	Participant Interviews	34
	Health Care Professionals Questionnaire	35
V.	Results and Preliminary Discussion	36
	A. Description of Sample by Diagnoses, Active Problems, and Self-Rated Health	36
	B. Dysfunction in Sample as Measured by the SIP	38

C. Age and Sex as Predictors of Health Status 41

D. Outcome of the Day Hospital Program as Measured by the SIP 42

E. Reliability of SIP Scores 45

F. Validity of SIP Scores 46

Relationship with Physician Diagnoses 48

Relationship with Active Problem List 52

Relationship with Self-Rated Health 52

Descriptive Validity 58

G. SIP Category Analysis 62

Independent Categories 64

Physical Dimension 69

Psychosocial Dimension 71

H. Participant Evaluation of the SIP 74

I. Evaluation of the SIP by Health Care Professionals 77

VI. Discussion 80

A. Descriptive Validity of the SIP 81

B. Construct Validity of the SIP 83

C. The SIP as an Outcome Measure of the Day Hospital Program 84

D. Age and Sex as Predictors of Perceived Health ..85

E. Concluding Remarks 86

REFERENCES 88

APPENDIX A: Letter of Introduction to Sample 107

APPENDIX B: Letter of Consent for Use of the Sickness Impact Profile 112

APPENDIX C: Participant Evaluation of the SIP 114

APPENDIX D: Health Care Professionals Questionnaire 116

APPENDIX E: Highly Prevalent Behavioural Dysfunctions of the Sample, as Measured by the SIP 118

APPENDIX F: Behavioural Dysfunctions Not Identified in Sample, as Measured by the SIP 120

APPENDIX G: Highly Prevalent Behavioural Dysfunctions of Depressed Patients, as Measured by the SIP 122

APPENDIX H: Behavioural Dysfunctions of All Patients with 'Poor' Self-Rated Health, as Measured by the SIP 124

List of Tables

Table		Page
IV.1	Description of Sample by Age and Sex	30
V.1	Self-Ratings of Health by Three Age Cohorts in Sample	38
V.2	Dysfunction in Sample as Measured by the Sickness Impact Profile (SIP)	39
V.3	Correlations Between Age, Sex, No. Diagnoses, No. Active Problems, Self-Rated Health, and Overall SIP	42
V.4	Paired t-tests for SIP Category, Dimension, and Overall Scores, T1 and T2	44
V.5	Internal Consistency of Dimension and Overall Scores of the SIP as Indicated by Cronbach's (1951) Alpha	46
V.6	Item/Category Correlations and Internal Consistency of SIP Categories	47
V.7	Comparison of Depressed Patients* with the Rest of the Sample** across Category, Dimension, and Overall SIP Scores	49
V.8	Highly Prevalent Behavioural Dysfunctions of Depressed Patients (n=15) as Measured in Categories EB, SI, and AB of the SIP	50
V.9	Highly Prevalent Behavioural Dysfunctions of Respondents Rating Their Health as Poor (n=6), as Measured by Category EB of the SIP	54
V.10	Overall SIP Scores by Self-Ratings of Health for Total Sample and Subgroups Defined on the Basis of Depression	56
V.11	Self-Ratings of Health by Depressed Patients (n=15) Compared with the Rest of the Sample	58
V.12	Interrelationships of SIP Categories	65
V.13	Interrelationships of SIP Independent Categories	66
V.14	Correlation Matrix of SIP Physical Dimension	70
V.15	Correlation Matrix of SIP Psychosocial Dimension	72

Table Page

V.16 Evaluation of the SIP by Respondents (N = 30).....75

V.17 Evaluation of the SIP by Health Care Professionals (N = 8).....78

List of Figures

Figure		Page
V.1	SIP Scores of Depressed Patients Compared to Rest of Sample	51
V.2	Self-Rated Health of Depressed Patients	57
V.3	Case 'A' Category Scores	60
V.4	Case 'B' Category Scores	61
V.5	Case 'C' Category Scores	63

I. Introduction

The aging of the Canadian population is well recognized, and this demographic trend is expected to continue for several decades (Stone & Fletcher, 1980). By the year 2030, it is estimated that the proportion of the Canadian population 65 and over will at least double, at which time it will represent 18% of the population of Canada (Statistics Canada, 1979).

As age is one of the prime determinants of the need for, demand for, and utilization of health care services (Bay, Leatt, & Stinson, 1982; Bay, Overton, Harrison, Stinson, & Hazlett, 1979), the increasing size of the older population has drawn attention to the importance of measuring the health status of the elderly (Canadian Governmental Report on Aging, 1982). In addition, wider awareness of the complexity of health measurement issues in terms of interactions between health, behaviour, and aging is occurring in the field of psychology (Birren, Cunningham, & Yamamoto, 1983; Siegler, Nowlin, & Blumenthal, 1980).

Linn (1976), among others, has asserted that "any assessment of the total functioning of an individual must take into account the subject's perception of his health or functional status" (p.119). This is a position which appears central to the study of health and behaviour among the aged. In recognition of the need for comprehensive, valid, and reliable measures of perceived health status suitable for use with the elderly, a clinical validation study of a

behaviourally based measure of perceived health status, the Sickness Impact Profile developed by Bergner et al. (1981), was undertaken.

In Chapter II, literature related to self-assessment of health is reviewed. Specifically, as self-ratings of health have been most commonly used as indicators of perceived health status among the elderly, a review of selected studies in the area is given. It is noted that a general impression conveyed by the literature is that self-ratings of health represent a "feeling-state" conception of health, to use Baumann's (1961) terminology, an orientation global in nature, largely inaccessible to external validation, and potentially influenced by social norms more than actual health status. Further, investigations of the sex morbidity ratio among the elderly based on self-ratings of health are confusing in their results.

In contrast, the relative advantages of a performance conception of health are examined. This is followed in Chapter III by a description of a behaviourally based measure of perceived health status, the SIP. Designed to operationalize perceived health in terms of function across a multidimensional spectrum of behaviours, the SIP appears to have considerable potential for application with elderly populations. Subsequent sections of this report present the methods used and results obtained in clinical validation of the SIP with patients attending a geriatric Day Hospital, followed by a discussion of the findings obtained.

II. Background to the Study

Given the importance of health as a variable in gerontological research (Abrahams, 1976; Elias, ELias, & Elias, 1977; Chappell, 1981), it is essential to consider that there is no simple or straightforward way to define or measure health. Health is a multidimensional concept (Kaplan, Bush, & Berry, 1979; George & Bearon, 1980; Chen & Bryant, 1975), with different dimensions suitable to different research questions and purposes. As stated by Sullivan (1966),

> Health is often spoken of as if it were a directly observable characteristic existing within an individual, but measurement of health, in fact, requires selection from many potentially measureable characteristics of a person or population. Another, equally important, consideration is the complex of circumstances surrounding the measurement process...There is little empirical justification for the assumption that a unidimensional continuum underlies and relates the measures referred to as healthy and unhealthy in different contexts. Health, defined without reference to a specific situation or purpose of measurement, may be merely a verbal artifact (p. 5-6).

Baumann (1961) recognized that health is, for many people, multidimensional, and she described three broad conceptions of health on which individuals may base appraisals of their own health status: a "feeling state orientation", a "symptom orientation", and a "performance orientation" (p. 40). Characteristic of responses classified as expressing a feeling-state orientation are phrases such as "I feel good today". Individuals who are symptom-oriented often describe the symptoms of their own diseases, and indicate that a healthy person would be one who did not have

these symptoms. Phrases such as "I can't get my breath" or "my legs give me a lot of pain" may be used by those who tend to focus upon symptoms. Among those who tend to view their health in terms of performance, the ability to carry on usual activities is central to their conception of health. Phrases such as "not doing heavy work", "use a cane to walk", or "make more mistakes than usual" may reflect this perspective.

A performance framework for measurement of health status appears to be particularly appropriate for use with elderly populations. Indeed, although health may be defined in physical, mental or social terms, health status of the elderly usually has one of two meanings: (1) the presence or absence of disease, or (2) degree of functional impairment (Stewart, Ware, & Brook, 1977; Shanas & Maddox, 1976; Chappell, 1981). If it is assumed that "those things that an old person says he can do are the best indicators of both how healthy he is and of the services he will require from the community", to cite Shanas (1974, p. 262), then *self-reports* of functional status are of key importance in measurement of the health of the aged.

Notably, however, previous research directed toward investigation of the self-assessed health of older persons has most often been based upon single item self-ratings of health rather than comprehensive self-reports of functional status. That is, evidence of Baumann's (1961) "feeling-state" orientation to health is more frequently

found in the literature than studies containing self-reports of the actual performance of activities.

The following overview of studies of self-rated health highlights previous theorizing about its nature and limitations, and provides a background for discussion of the relative advantages of behaviourally based measures of perceived health status.

A. Selected Studies of Self-Rated Health

There is no clear consensus concerning the meaning, and more importantly, the utility of one of the most frequently assessed constructs in gerontological research, that of self-rated health. Typically, respondents are asked the question, "Compared to others your age, do you consider your health to be excellent, good, fair, or poor?". Despite widespread usage and extensive research directed toward identifying its nature and determinants (eg. Wan, 1976; Graney & Zimmerman, 1980-81), self-rated health remains a vague concept inaccessible to external validation. Nevertheless, it has been linked to a variety of constructs, as indicated by an overview of studies in the area.

Suchman, Phillips, & Streib (1958) proposed that self-ratings of health show significant relationships to both attitudinal and behavioural correlates of health, and that they measure something different from physician's ratings, what they called "perceived" or "subjective" health as opposed to "actual" or "objective" health. Subsequently,

Friedsam and Martin (1963) argued that what was at issue was not a question of "subjective" or "perceived" health as against "objective" or "actual" health, but rather different dimensions of health. It was also their suggestion "that when the relationships have been more fully explored, the self-rating of health may prove to be, particularly for older people, a rather simple and sensitive barometer of self-image" (p. 183).

Maddox (1962 & 1964) proposed that pessimism about one's health, as reflected in overestimation of poor health, is associated with a general anxiety about health and poor adjustment to the environment, while optimism about one's health is associated with the absence of these two characteristics. He also viewed self-ratings of health as indicators of adoption or rejection of the sick role. In a follow-up study of the subjects involved in the 1962 and 1964 research, Maddox and Douglass (1973) reported "a persistent, positive congruence of self and physician's ratings of health", and concluded older people "have and maintain a strong reality orientation" with regard to perceived health status (p. 92).

Noting the previous work of Suchman et al. (1958), Friedsam & Martin (1963), and Maddox (1962), among others, Tissue (1972) sought to identify those factors which are most closely associated with self-rated health in old age. In a sample of 256 low income elderly, self-ratings of health were observed to be most strongly related to other

subjective indicators of health (ie. incidence of health-related worry, perception of recent changes in health, feeling 'old and tired', and pain). That is, Tissue (1972) found general self-ratings of health to reflect the "way one evaluates or perceives one's individual condition to a greater extent than it does functional capacity, actual incidence of disease, or receipt of medical care" (p. 93).

In a review of thirty years of research on the subjective well-being of older Americans, Larson (1978) discovered that, among all the elements of an older person's life situation, self-rated health was most strongly related to subjective well-being. In addition, it was the most important factor influencing individual assessment of life satisfaction as evidenced, for example, in studies by Palmore & Luikart (1972), Edwards & Klemmack (1973), and Spreitzer & Snyder (1974).

Similarly, Snider (1980) found support for the trend in the literature that of the several variables reported to be the most reflective of life satisfaction among the elderly, self-rated health was specifically the most dominant on the basis of presence and strength of association as well as absolute and relative proportion of the total variance explained.

In a similar vein, Mancini & Quinn (1981) investigated the interplay between self-ratings of health and health behaviour in terms of their importance for morale in old age. Respondents in their sample (N=104) had a mean age of

72.5 years, and about 67% reported themselves as either in "good" or "excellent" health. In addition, those who generally saw themselves in good health had higher morale.

Mossey and Shapiro (1982) used data from the Manitoba Longitudinal Study on Aging to test the hypothesis that self-rated health is a predictor of mortality independent of objective health status. Subjects were a random sample of non-institutionalized residents of Manitoba aged 65+ in 1971 (N=3,128). Self-rated health was defined by responses to the question, "For your age would you say, in general, your health is excellent, good, fair, poor, or bad?" Analysis of the data revealed that, controlling for objective health status, age, sex, life satisfaction, income, and urban/rural residence, the risk of early (1971-1973) and late mortality (1974-77) for persons whose self-rated health was poor was 2.92 and 2.77 times that for those who rated their health as excellent.

Interpretations suggested by these researchers included the possibility that self-rated health is a sensitive indicator of physical well-being, that persons with a durable advantage derived from protective health habits are the ones who report good health for any current level of objective health, and that emotional health, including positive feelings, could be reflected in self-rated health and its correlation with longevity. In any event, Mossey & Shapiro (1982) concluded that their findings provide "empirical support for the long held, but inadequately

substantiated, belief that the way a person views his health is importantly related to subsequent health outcomes" (p. 800).

To summarize, self-ratings of health have been variously described as measuring "perceived" versus "actual" health (Suchman et al., 1958), self-image (Friedsam & Martin, 1963), anxiety and adjustment to the environment (Maddox, 1964), adoption or rejection of the sick role (Maddox, 1962 & 1964), subjective well-being (Larson, 1978), life satisfaction (Snider, 1980), morale (Mancini & Quinn, 1981), physical well-being (Mossey & Shapiro, 1982), and emotional health, including "positive feelings" (Mossey & Shapiro, 1982). As such, self-ratings of health are primarily indicative of "feeling states", and they require the individual to make an implicit judgment based upon whatever his or her concept of health happens to be (Parkerson, Gehlbach, Wagner, James, Clapp, & Muhlbaier, 1981). In the absence of a commonly accepted standard of what is excellent, good, fair, or poor health, self-ratings of health "may reflect the norms of a particular group relevant to the respondent, for instance, a group based on sex, age, or residence", to quote Fillenbaum (1979, p. 45). In this event, self-ratings of health may be influenced more by social norms than actual health status.

Indeed, review of the literature related to health and behaviour among the aged suggests social norms linked to sexual status may contribute to the sex morbidity ratio.

Given that this ratio is based upon survey research in which respondents report whether or not they have, for example, a specific type of health problem, it has been hypothesized that the excess morbidity of elderly women may be due to their greater willingness to report such data (Payne & Whittington, 1976). Further, if a positive relationship between reporting behaviour and perceived health status is assumed, the question arises as to whether elderly men and women tend not to perceive their health in the same way. In this light, the next section of summarizes current epidemiological data on the sex morbidity ratio among the elderly, and examines the interpretation that elderly women and men tend to assess their health differently.

B. The Sex Morbidity Ratio

The Canadian Governmental Report on Aging (1982) noted that in 1976 the expected remaining years of life for men at age 65 were 13.95 years, and for women, 18 years.

In the fifty years between 1921 and 1971, life expectancy at the older ages increased much more sharply among females than males in this country. For males aged 65 the increase was about one half a year (13.2 to 13.7), whereas for females the corresponding data were 13.7 to 17.5 (Stone & Fletcher, 1980).

The excess mortality of men is often compared to the excess morbidity for women. In terms of the sex differential for morbidity, the Canada Health Survey (1981) found in the

65 and over age group that females have more health problems than males. Specifically, elderly women reported mental disorders, diabetes, thyroid disorders, anemia, headaches, sight disorders, hypertension, heart disease, acute respiratory illness, influenza, hayfever and other allergies, dental problems, digestive disorders, skin disorders, arthritis and rheumatism, limb and joint disorders, trauma, and other health problems more frequently than did elderly men. In contrast, males 65 and over reported more hearing disorders, bronchitis and emphysema, asthma, and gastric and duodenal ulcers than did their female cohorts (Table 57, Prevalence of Health Problems by Age and Sex, by Type of Health Problem, Canada, 1978-79, p. 115).

Females aged 65 and over also have a higher number of annual bed-days per person (15.28 vs. 10.55), and a higher number of annual major activity loss days per person (15.59 vs. 10.13) than do males 65+ (Canada Health Survey, 1981, p. 119 & 120). An earlier Canadian survey by public health nurses of the health and living conditions of the aged found as well that women reported more symptoms and illness and had more bed-sickness than men (Schwenger & Sayers, 1971). However, Stone & Fletcher (1980) have observed that the pattern of differences in male/female hospital utilization among the 65 and over age group varies sharply from that of the whole population. Females in the whole population have much higher rates of hospital utilization than males,

whereas elderly females have much lower rates than elderly males.

In terms of functional impairment, the Framingham Disability Studies of physical and social disability among the aging (n=2,654) indicated women had more unmet housekeeping and transportation needs than men (Jette & Branch, 1981; Branch & Jette, 1981). This difference was significant within each of three age cohorts (55-64 years, 65-74 years, & 75-84 years). Women were also more disabled than men in activities of daily living such as bathing, dressing, and transferring from one position to another. In addition, there was no support for the interpretation that the greater prevalence of physical disability among elderly women was an artifact of the increased proportions of women living into old, old age.

Thus, in contrast to the excess mortality of elderly males, elderly females seem disadvantaged in terms of excess morbidity. This appears to hold true regardless of whether health is measured by the presence or absence of disease or in functional terms. Various interpretations of this phenomenon are examined in the next section, with particular attention directed to the hypothesis that elderly men and women tend to assess their health differently.

Interpretations

A general impression conveyed in the literature is that women are more willing to report symptoms of distress and

illness than are men (Verbrugge, 1979). That is, one resolution of the paradox of women's greater longevity and poorer health rests in the argument that women are not less healthy than men but do have a greater propensity to underrate their health and to seek medical care (Payne & Whittington, 1976). For example, Verbrugge (1976a) has stated, "We suspect that males are physically sicker from both acute and chronic conditions but that social and psychological factors enhance females' reports to such an extent that a female excess appears overall" (p.399). Similarly, Mechanic (1978) noted that women "report many more subjective symptoms than men, particularly those symptoms that may be indicative of physical or psychological distress". And, he concludes, "it is difficult to support the argument that women have more illness than men" (p.213).

Dohrenwend and Dohrenwend (1977) have also argued that the higher rates of morbidity in women found in community surveys of mental health are a consequence of women being more willing to engage in illness behaviour than men. Such help-seeking patterns are said to reflect, in turn, social norms that permit women to complain more readily and to appear less stoic; the corollary assumption holds that men often overrate their health or ignore health problems in an attempt to fulfill a sociocultural expectation that men ought to be strong, healthy, and stoical in the face of pain (Nathanson, 1975; Verbrugge, 1976a, & 1976b; Mechanic, 1978; Phillips & Segal, 1969; Maddox, 1964).

Typical of the "sex role socialization" interpretation is the following quotation from Shanas (1968):

> In Denmark, Britain, and the United States ... men are the dominant sex. They are expected to be stronger than women, hardier, and less complaining. This cultural pattern is followed throughout the lifespan and into old age. Consequently, old men in comparison with old women are more likely to say their health is "good" (p.214).

Yet, conflicting evidence does exist for the assumption that elderly women tend to underrate their health and elderly men to overrate their health. Ferarro (1980) found that although elderly males in his study generally reported less disability and fewer physiological disorders than the elderly females, these same men also rated their health as poorer than the women. He states, "In comparison with females ... males tend to rate themselves as having poorer health *even though* they report less disability and fewer illnesses" (p. 381).

Fillenbaum (1979) comes to a similar conclusion in her interpretation of the finding that for any given self-rating of health on the scale "excellent", "good", "fair", or "poor", women overrated their health in relation to their objective health status. That is, women whose health tended to be objectively poorer as indicated by the number of health-related problems reported present in the previous month, the number of different types of medications taken during the previous month, and the number of different illnesses and disabilities presently affecting each respondent, tolerated more health problems for a given

self-rating of health than did their male counterparts.

Further, Chappell & Havens (1980) report that "if the conventional age 75 is chosen, old elderly women emerge as expressing greater overall well-being than old elderly men" (p.162). This, they suggest, lends support to Dulude's (1978) contention that women do not translate objectively worse situations to the psychological/subjective level. Although the validity of their position is uncertain, Chappell & Havens (1980) speculate that "a woman's role throughout her lifetime is characterized by change and inconsistency, which leads to the resilience and strength of older women" (p. 163), a view in keeping with that of Beeson (1975), Kline (1975), and Giesen & Datan (1980).

Unfortunately, an apparent weakness of the studies by Ferraro (1980), Fillenbaum (1979), and Chappell & Havens (1980) is that their supposed indicators of objective health status were actually based on the self-reports of subjects rather than on clinical measures independent of reporting behaviour. Thus, it is noteworthy that Davis (1981) found no support for the hypothesis that women are more willing to over or underreport symptoms in a study incorporating measures of disease independent of symptom reports. Reporting of chronic joint symptoms for both men and women was significantly related to radiographic evidence of osteoarthrosis and two measures of treatment behaviour.

Atchley (1976) cast a further shadow upon previous theorizing about the social and psychological differences

between men and women in later life, including how they perceive their health, in his observation that studies which have included sex comparisons usually lump all respondents into a single age category, 65 and over. He asserts, "Even those studies that do include more detailed age breakdowns never control for the sex differences in marital status, education, and income", and "because these comparisons failed to control alternative causes of sex differences they are of limited value" (p. 204). Notably, his empirical study of male/female comparisons in late life found self-rated health was not significantly different when age, marital status, education, and income adequacy were controlled.

Controlling for physical health in each case, Blazer & Houpt (1979) conducted a survey of the perception of poor health in the healthy older adult (N=977). Contrary to previous data (e.g. Maddox, 1962; Tissue, 1972; Maddox & Douglass, 1973; Friedsam & Martin, 1963), perceived poor health in the healthy elderly was not associated with age or sex. Rather, the physically healthy elderly who perceived their physical health status to be poor were more depressed, more hypochondriacal, and dissatisfied with life. They tended to complain of multiple symptoms, activities of daily living were decreased, and they were more likely to visit their doctor frequently during the year. That is, the presence of depression, hypochondriasis, and dissatisfaction with life were more important predictors of self-rated health than the sex or age of the subjects.

In summary, previous theorizing and empirical research directed toward interpretation of the sex morbidity ratio in the elderly is confusing in its conclusions. There is evidence for underrating and overrating of health by both sexes (e.g. Shanas, 1968; Ferraro, 1980; Fillenbaum, 1979) while nonsignificant differences beween the sexes have also been reported (e.g. Atchley, 1976; Blazer & Houpt, 1979). Further, mental depression, hypochondriasis, and dissatisfaction with life may have greater impacts upon perceived health status, as indicated by self-ratings of health, than actual health status does.

In order to establish more clearly whether or not the sex morbidity ratio is partly attributable to a tendency for men to view their health differently from women, it would seem preferable to operationalize perceived health within a performance framework. Also, it is suggested that should estimates of the actual performance of activities be substituted for self-ratings of health, the sex morbidity ratio among the elderly may be more clearly interpretable in terms of existent differences in physical and psychosocial health status.

C. The Performance Conception of Health

The advantages of definitions of health in terms of performance or function have been recognized for some time. Patrick, Bush, & Chen (1973), for example, considered such definitions of health as essential for bridging the gap

between theoretical concepts of health and their empirical measurement. Alternatively, Reynolds, Rushing, & Miles (1973) asserted that health should be defined in terms of behaviour "in view of the problems of reliability and validity associated with nonbehaviour measures, such as perceived health and self-diagnosis", and as "it is in terms of behavioural functioning rather than perceived health ... that the concept of health is most relevant for social system functioning" (p. 272).

Bergner et al. (1976a) have also proposed that a performance conception of health has several attributes which make it a useful basis for a measure of sickness-related behaviour. With reference to earlier work by Fabrega (1975), they state,

> First, performance or behaviour may be reported directly by the individual under consideration. Second, performance may also be observed and reported by another respondent about the individual under consideration. Third, performance may be affected by medical treatment even though the disease itself may be unaffected. Fourth, performance can be measured whether or not the individual is receiving medical care. And fifth, a measure based on performance or behaviour permits relating diverse definitions of disease and sickness by uncovering universal patterns of behavioural dysfunctions (p. 396-397).

A performance conception of health also permits measurement of perceived health across a multidimensional spectrum of areas of living, and in so doing facilitates collection of qualitative detail regarding perceived strengths and weaknesses in functional status. This information, in turn, may be used as an outcome measure of a

health care program from the perspective of of the patient (Brook, Davis-Avery, Greenfield, et al. 1977). In addition, given the tendency to chronic illnesses and functional disability among the elderly (Shanas & Maddox, 1976), a performance conception of health appears preferable to the more customary self-ratings in any study involving interactions between health, behaviour, and aging.

In this regard, the Sickness Impact Profile (SIP) is a behaviourally based measure of perceived health status developed by the Department of Health Services, University of Washington, Seattle (Bergner et al., 1981; Pollard et al., 1978; Bergner et al., 1976a; Bergner et al., 1976b; Pollard et al., 1976; Carter et al., 1976; Gilson et al., 1975). The SIP focusses upon whether an individual perceives himself as actually performing an activity as it is described in a SIP statement. That is, it relies exclusively upon the self-assessment of the individual within a performance framework for measurement of health status.

Although the SIP was not specifically designed for use with the elderly, its performance base and and comprehensive nature indicated that it could be extended for use with the 65+ population. Thus, the main purpose of this study was clinical validation of the SIP for measurement of the perceived health status of a small group of elderly people attending a geriatric Day Hospital.

More specifically, the objectives were:

1. To assess the descriptive validity of the Sickness Impact Profile (SIP) as a measure of the perceived health status of a small group of elderly people attending a geriatric Day Hospital.

2. To assess the construct validity of the SIP in terms of its relationships with indicators of objective health status and patient self-ratings of health.

3. To assess the utility of the SIP as an outcome measure of the Day Hospital program.

4. To examine the relationships between age, sex, objective health status, and self-assessed health of the participants in the study.

Several terms used in the study required specification as in many cases there were restrictions in common usage.

Chronic disease: all impairments or deviations from normal which have one or more of the following characteristics: are permanent; leave residual disability; are caused by nonreversible pathological alteration; require special training of the patient for rehabilitation; may be expected to require a long period of supervision, observation, or care. (U. S. Commission on Chronic Illness, 1957, cited in Mausner & Bahn (1973)).

Elderly: a person who is 65 years of age or older.

Functional disability: functional limitations in the performance of normal daily role activities as a result of illness or injury (Stewart et al., 1977).

Health: a dynamic continuum of physical and psychosocial functional states.

Reliability: refers to the consistency or stability with which an index measures a construct. An index with a low reliability does not possess the required sensitivity or precision to differentiate individuals or groups in terms of health status unless the differences in health status are of the grossest nature (Chen & Bryant, 1975).

Sickness: an individual's own perception of his health.

Sickness impact: self-reports of dysfunction, clinical reports of dysfunction, others' reports of dysfunction, and tests of dysfunction using other instruments (Bergner et al., 1981).

Validity: the range of inferences that are appropriate when interpreting a measurement, a score, or the result of a test (Cronbach, 1971). That is, the validity of a measure defines the meaning of a score derived from it. Validity is not

absolute; it is relative to the domain about which statements are made. An indicator or index is a valid measure of health status only to the extent that it expresses or quantifies that construct (Kaplan, Bush, & Berry, 1976).

In the next section a description of the Sickness Impact Profile is provided, with particular attention directed to its purpose, development, scaling, scoring, reliability, and validity.

III. The Sickness Impact Profile

Purpose

The SIP was developed to provide a measure of perceived health status that is sensitive enough to detect changes or differences in health status in a variety of circumstances. It is considered to be broadly applicable across demographic and cultural subgroups, and may be used at the level of (1) the individual patient, (2) small groups, and (3) general populations. Further, it is intended "to assess a wide range of sickness-related impacts varying from minor to very severe dysfunctions" (Bergner & Gilson, 1981, p. 140), and "to provide a measure of the effects or outcomes of health care that can be used for evaluation, program planning and policy formulation" (Bergner et al., 1981, p.787).

The SIP contains 136 items each describing a sickness-related behavioural change. The items are grouped into 12 categories or areas of living: Sleep and Rest (SR), Emotional Behaviour (EB), Body Care and Movement (BCM), Home Management (HM), Mobility (M), Social Interaction (SI), Ambulation (A), Alertness Behaviour (AB), Communication (C), Work (W), Recreation and Pastimes (RP), and Eating (E). Categories EB, SI, AB, and C are aggregated into a "psychosocial dimension", and categories BCM, M, and A are aggregated into a "physical dimension". The five remaining categories are not aggregated in any way.

Development, Scaling, and Scoring of the SIP

As stated by Bergner et al. (1981), "The strategy chosen for developing, assessing, and revising the SIP was based on methodological principles that emphasized the evaluation of reliability and validity in a variety of settings, the determination of the relationship of the SIP to other measures currently in use and the evaluation of its unique contribution as an outcome measure of health status" (p. 788).

Development of the SIP began in 1972 with the initiation of procedures to collect and evaluate statements describing sickness-related behavioural dysfunction. Statements were gathered from patients, individuals caring for patients, the apparently healthy, and health care professionals. In addition, the scientific literature on health measurement was reviewed and additional statements were gathered from previously developed instruments containing references to behavioural dysfunctions (Bergner & Gilson, 1981). A total of 1250 statements were screened to yield 312 items grouped into areas of activity or categories and then included in a prototype Sickness Impact Profile. This questionnaire was the subject of field trials conducted in 1973, 1974, and 1976 (Bergner et al., 1981).

Sampling during the field trials was designed "so that the instrument would be tested on subjects that spanned a range of type and severity of illess" (Bergner et al., 1981, p.788), and included home care patients, walk-in clinic

patients, rehabilitation medicine outpatients and inpatients, speech pathology inpatients, outpatients with chronic health problems, and a group of enrollees in a pre-paid health plan.

During development of the instrument, a total of 133 judges rated the relative severity of dysfunction of each statement so that scale values could be assigned (Carter, Bobbitt, & Bergner, 1976). The judging group was composed of 108 health consumers and 25 health care professionals and pre-professional students. They were asked to rate each statement in terms of severity without regard for who might be experiencing it, what might be causing it, or future implications (Bergner & Gilson, 1981). Agreement among the judges was very high (Carter et al., 1976).

In scoring the SIP, scale values are used to calculate category, dimension, and overall SIP scores. Possible scores in each category range from 0 to 100, and higher scores reflect worse function. A score for the overall SIP instrument is computed by summing the scale values of all items checked and dividing that sum by the grand total of all statement scale values. This ratio is then multiplied by 100 to convert it to a percent. Similarly, scores for each category are calculated.

Reliability

Test-retest reliability (ie. reproducibility) and internal consistency of the SIP were comparable across all

field trials. In the 1976 field study, test-retest reliability was 0.92 for overall SIP scores and 0.50 for category items, while internal consistency as indicated by Cronbach's alpha (Cronbach, 1951) was 0.94 (Bergner et al., 1981).

The 1976 field trial compared the reliability of three types of administration of the SIP: interviewer administration (I), interviewer-delivered self-administration (ID), and mail-delivered self-administration (MD). High levels of test-retest reliability, evaluated in terms of score correlations, were demonstrated for the I's (r=0.97) and the ID's (r=0.87). No retests on MD's could be obtained (Pollard et al., 1976). Internal consistency measured in terms of Cronbach's alpha (Cronbach, 1951) was 0.94 for both I's and ID's but only 0.81 for MD's (Bergner et al., 1981).

Validity

The construct validity of an instrument is the degree to which it measures the theoretical construct or trait that it is designed to measure (Allen & Yen, 1979).

In this regard, the validity of the SIP was tested during field trials by observing how strongly it correlated with other instruments purporting to measure health status or some specific aspects of it. Two further validation efforts involved comparing the SIP with judgments of observers assessing roughly the same construct, specifically

the patient's judgment of his health and sickness, and the clinician's assessment of the patient's health and dysfunction (Bergner et al., 1976a).

Using 1976 data (N=230), the correlation between SIP scores and self-assessment of dysfunction was 0.69; between SIP scores and self-assessment of sickness was 0.63; between SIP scores and selected questions from the National Health Interview Survey Index of activity limitation was 0.55; between SIP scores and clinician assessments of dysfunction was 0.50; and between SIP scores and clinician assessments of sickness was 0.40 (Bergner et al., 1981; Bergner & Gilson, 1981; Bergner et al., 1976a). Multitrait-multimethod techniques, as developed by Campbell and Fiske (1959), and multiple regression techniques further suggested that "SIP scores are more highly related to those criterion measures that were, a priori, considered to be most reflective of the construct of sickness and the methodology employed in the SIP" (Bergner et al., 1981, p. 797).

The validity of the SIP was also investigated through assessment of its relationship to objective clinical data used to follow the progress of patients with specific diagnostic conditions. That is, the "clinical validity" of the SIP was determined for patients in the disease categories of total hip replacement, hyperthyroidism, and rheumatoid arthritis. To cite Bergner and Gilson (1981), "As a general statement it may be said that the SIP showed a high correlation with the established and accepted tests in

all three disease conditions and at all of the stages of the diseases that were tested" (p. 144).

The descriptive validity of the SIP, or "the extent to which dimension and category scores and item-checking patterns provide a useful and meaningful qualitative description of different samples and types of subjects", was assessed by pattern and profile analyses of individual and diagnostic group data (Bergner et al., 1981, p. 799). Findings were described as consistent with clinical observations, while providing information that in some cases was new and in others highlighted clinical data that may have been ignored or de-emphasized. Further, characteristic patterns of SIP category and dimension scores across all hip patients and across all hyperthyroid patients permitted differentiation among diseases and among patients (Bergner & Gilson, 1981).

Based on the research that has been done, the Sickness Impact Profile appeared to have considerable potential for application with the elderly. Therefore, a clinical validation study of the SIP with a group of elderly patients was proposed and subsequently conducted in a geriatric Day Hospital affiliated with the Department of Geriatric Medicine, University of Alberta. A description of the methods utilized in this research is provided in the next section.

IV. Methods

A. The Sample

Thirty people attending a geriatric Day Hospital located in the city of Edmonton volunteered to participate in the study. The sample was obtained by inviting each person admitted to the Day Hospital between January 1, 1983 and March 31, 1983 to participate, with the exception of those who were discharged soon after admission, were mentally confused, or unable to read or understand spoken English. A letter of introduction described the nature and purposes of the study to each potential respondent. A copy of this letter is located in Appendix A. Written consents, witnessed by a staff member of the Day Hospital, were obtained for each member of the sample.

Descriptive data for the sample are summarized in Table IV.1. Mean age was 74.7 years, with a range of 65 to 93 years. Thirteen people were 65 to 74 years of age, and 17 people were 75 years of age or older. There were 13 males and 17 females.

All participants in the study resided in the community and attended the Day Hospital once or twice weekly. In keeping with the criteria for admission to Day Hospital, they required assessment and/or the provision of integrated medical, nursing, rehabilitation, and social services. Specific services accessed by the respondents included medical assessment, care and supervision, nursing care and

TABLE IV.1
DESCRIPTION OF SAMPLE BY AGE AND SEX

AGE (YRS)	SEX		TOTAL
	Male	Female	
65-74	5	8	13
75+	8	9	17
TOTAL	13	17	N=30

supervision, diagnostic testing, physiotherapy, occupational therapy, recreational and diversional therapy, socialization groups, speech therapy, social services, individual and family counselling, medication monitoring and counselling, diet therapy and counselling, and transportation to and from the program.

B. Research Design

Participants were assessed on two occasions using measures of perceived health status as indicated by self-ratings of health and responses to the Sickness Impact Profile, and of objective health status as indicated by the active problems listed in each respondent's medical record. Approximately one month elapsed between the first and second occasions.

A *pilot study* based on the first five members of the sample resulted in revisions to the measure proposed to summarize objective health status. Rather than constructing

an index of objective health by adding the number of different diagnoses and the number of active problems listed, these two variables were regarded as independent indicators of objective health status.

C. Instruments

The data collection instruments were of three types listed as follows. Each is described separately.

1. Indicators of Objective Health Status

   1.1 Number and types of diagnoses made by the physician of the Day Hospital;

   1.2 Number and type of active problems listed in problem oriented medical records.

2. Indicators of Self-Assessed Health

   2.1 Self-ratings of health;

   2.2 The Sickness Impact Profile (SIP).

3. Evaluations of the SIP

   3.1 Participant Interviews;

   3.2 Health Care Professionals Questionnaire.

Physician Diagnoses

Upon admission to the Day Hospital, all patients undergo a medical examination by the same physician. Based upon this assessment, the diagnoses made for each member of

the sample were listed by this physician and in the present study were used as an indicator of objective health status.

Active Problems Listed in Problem Oriented Medical Records

The Day Hospital utilizes the problem oriented medical record (POMR) system of record keeping which consists of four main parts: a problem list, a defined data base, plans, and progress notes. The problem list is considered to be the key to this system of medical records insofar as it is a "list of the clinically significant events in a patient's life and highlights the factors which might affect management" (McIntyre, 1979, p.39).

A list of problems is identified for every Day Hospital patient following a multidisciplinary assessment process. Each problem is validated by the patient prior to inclusion in his medical record. Problems which require active intervention are called "active" problems, in contrast to those identified as "inactive" (that is, not requiring immediate action for the purposes of diagnosis, monitoring, treatment, or patient education), or "resolved".

The number and type of active problems listed for each participant were recorded on the same day as the first and second interviews.

Self-Ratings of Health

Self-ratings of health were obtained in answer to the question, "Compared to others your age, how do you rate your

health?", given the response categories "Excellent" (1), "Good" (2), "Fair" (3), or "Poor" (4).

The Sickness Impact Profile

The SIP was described in detail in Chapter III. As interviewer administration of the SIP yielded the most reliable data during the 1976 field trial, this was the type of administration selected. All SIP's were administered by this investigator according to procedures recommended by Conn, Bobbitt, and Bergner (1978). It was emphasized to respondents that they should respond to (check) *only* those statements in the SIP that they were *sure* described them today and were related to their state of health. Certain individuals required assistance to read SIP statements because of poor vision or illiteracy. For those with hearing problems, it was necessary to read each item in a loud voice.

Test-retest reliability estimates and the internal consistency of category, dimension, and overall SIP scores were calculated in this study to permit comparison with previous field trials. These estimates also assisted in assessment of the validity of the SIP scores obtained as descriptors of the perceived health status of the study sample.

The validity of the SIP was assessed by observing its relationship to physician diagnoses, active problems listed in problem oriented medical records, and self-ratings of

health. In addition, participants in the study and health care professionals of the Day Hospital were asked to give their opinions about the SIP. Particular attention was directed toward evaluation of the SIP's ability to discriminate between subgroups within the sample, and to provide individual patient profiles consistent with clinical reports.

A copy of the letter from Dr. M. Bergner in which permission was granted for use of the Sickness Impact Profile is in Appendix B.

Participant Interviews

Immediately following the second administration of the Sickness Impact Profile, the interviewer posed three questions to each respondent. Firstly, he was asked to indicate whether or not the SIP gave an accurate picture of how he saw his health at this time. Secondly, the respondent was asked if he had difficulty in understanding the instructions for responding to the questionnaire. Thirdly, the respondent was asked, "Do you think that the information you provided in answering the Sickness Impact Profile should be shared with the health professionals directly concerned with your program at the Day Hospital?".

The first and second questions aimed to elicit participants' opinions about the strengths and weaknesses of the SIP as a tool for describing how they saw their health. The third question was posed to determine the importance

attributed by a respondent to his SIP data, and to ascertain whether he was willing to release this information to the health professionals directly involved with his program at the Day Hospital.

Appendix C contains a copy of the questionnaire used in the participant interviews.

Health Care Professionals Questionnaire

The health care professionals of the Day Hospital familiar with the Sickness Impact Profile were requested to evaluate it according to seven criteria. This evaluation was conducted on the last day of the data collection phase, and subsequent to two presentations in which the purpose, format, content, and information obtained by the SIP were described and illustrated with data from the sample.

Appendix D contains a copy of the questionnaire completed.

V. Results and Preliminary Discussion

The initial three sections of this chapter are essentially descriptive in nature. Firstly, the sample is described in terms of the most frequently diagnosed disease states, the modal number of active problems listed, and self-ratings of health. Secondly, dysfunction in the sample as measured by the SIP is presented. Thirdly, the absence of a significant relationship between the age and sex of the respondents and indicators of objective and subjective health status is discussed in terms of sample characteristics.

The focus of the remaining sections is upon evaluation of the SIP with regard to its utility as an outcome measure of the Day Hospital program, as well as its reliability, validity, and internal structure as indicated by category analysis. Lastly, evaluations of the SIP by participants in the study sample and the health care professionals of the Day Hospital are reviewed.

A. Description of Sample by Diagnoses, Active Problems, and Self-Rated Health

The predominance of chronic disease states among the sample was evident in the nature of the diagnoses reported. The most frequently diagnosed conditions were osteoarthritis, depression, hypertension, chronic obstructive pulmonary disease, and cataracts. A tendency to multiple health problems was indicated by a mean of 4.37

diagnoses per respondent, with a range of one to eight.

Initially, the modal number of active problems listed in the problem oriented medical records of the participants in the study was 9 (range four to eighteen problems). However, the mode decreased to 8 (range one to 15) by the time of the second participant interviews.

With regard to self-ratings of health, 13 people said their health was "good", 11 said it was "fair", and 6 said it was "poor". No one selected the response "excellent".

The relationship between self-ratings of health and age was not significant when the sample was dichotomized into two subgroups (65 to 74 years, and 75 years and older). However, a tendency for self-ratings of health to be more positive among the oldest members of the sample was evident when the sample was divided into three subgroups (65 to 74, 75 to 84, and 85 years of age or older). With reference to Table V.1, the observation that all respondents in the oldest age range stated their health was "good" is consistent with earlier reports by Shanas et al. (1968), and Ferraro (1980), among others.

Shanas et al. (1968) found in cross-cultural research that "there is no marked decline with age in the proportion of old people who feel that their health is good" (p.214). Similarly, in a study of low income elderly (N=3,370) Ferraro (1980) discovered that though members of his sample 75 years of age or older were more disabled and reported more physiological disorders, they also tended to express

TABLE V.1
SELF-RATINGS OF HEALTH BY THREE AGE COHORTS IN SAMPLE

AGE (YRS.)	POOR	FAIR	GOOD	EXC.	TOTAL
65-74	5	6	2	0	13
75-84	1	5	6	0	12
85+	0	0	5	0	5
TOTAL	0	13	11	6	N=30

better health than those in the age range 65-74 years when analyses controlled for other pertinent factors. His observation that "the oldest of the elderly population are not as likely to reflect negative views of their own health" and that "inspite of the difficulties that they experience due to physiological disorders they tend to express a more positive view of their own health" (p.380) appear substantiated in this study. That is, although the participants in this study 85 years of age or older (n=5) generally had higher overall SIP scores, they all said their health was "good".

B. Dysfunction in Sample as Measured by the SIP

The structure of the SIP and mean scores in each category for the entire sample are shown in Table V.2. For comparison, data from a previously studied general population sample (Gilson, Bergner, Bobbitt et al., as cited in Deyo, Inui, Leininger, & Overman, 1983) are included.

TABLE V.2
DYSFUNCTION IN SAMPLE AS MEASURED BY THE SICKNESS IMPACT PROFILE (SIP)

CATEGORY OR DIMENSION	ITEMS IN SCALE, NO.	PROFILE SCORE (n=30) MEAN	(S.D.)	PROFILE SCORE (n=624) FOR GENERAL POPULATION* MEAN	(S.D.)
Overall SIP	136	19.7	(11.0)	3.6	(5.3)
Dimension I: Physical	45	20.8	(12.7)		
Ambulation	12	26.3	(16.7)	3.1	(7.0)
Body Care And Movement	23	17.9	(15.5)	1.0	
Mobility	10	22.4	(14.5)	2.7	(7.3)
Dimension II: Psychosocial	48	19.8	(15.7)		
Emotional Behavior	9	25.9	(26.4)	3.8	(9.6)
Social Interaction	20	16.4	(10.6)	5.2	
Alertness Behavior	10	26.9	(30.6)	4.0	(8.7)
Communication	9	13.2	(14.5)	1.1	(3.6)
Independent Categories					
Sleep And Rest	7	21.0	(21.3)	7.2	(13.2)
Household Management	10	31.1	(24.5)	5.4	(12.5)
Work	9	0.5	(2.7)	8.5	(19.4)
Recreation And Pastimes	8	33.1	(25.8)	10.2	(15.8)
Eating	9	6.7	(5.1)	1.6	(3.3)

* A stratified random sample from a previous study in a large U.S. prepaid health plan. Men and women were equally represented as were three age strata (18 to 44, 45 to 64, and 65 to 74 years). The SIP was self-administered after verbal instructions (Gilson, Bergner, Bobbitt, et al, 1979; cited in Deyo et al, 1983).

A list of the behavioural dysfunctions identified by 50% or more of the sample is provided in Appendix E. The most highly prevalent dysfunctions reported were in terms of sleep and rest, special diets, doing heavy work at home, walking, and going out in the community for social visits and entertainment.

Upon review of mean SIP category scores, the three highest levels of dysfunction in the sample were in the areas of recreation and pastimes ($\bar{X}$=33.1), home management ($\bar{X}$=31.1), and alertness behaviour ($\bar{X}$=26.9). It was interesting to note that the items most frequently checked tended to be the lowest weighted items in each category. That is, on a continuum from minimal to maximal severity of dysfunction, this sample generally occupied the lower end of the scale. Given that the range of possible scores in each category is 0 to 100, with a higher score indicating worse function, most people had scores less than 50 in every category. Indeed, a floor effect was evident insofar as scores of 0.0 were obtained by most respondents (76.7%) in at least two categories.

Twenty items, primarily in the categories "Work" and "Eating", not checked by any member of the sample are listed in Appendix F. These items were generally not applicable as all but one person was retired, and most had no problems related to eating except for the need for a special diet.

Nevertheless, in comparison to the sample (n=624) used in the 1976 field trial of the SIP, the participants in this

study had much worse health.

C. Age and Sex as Predictors of Health Status

Contrary to findings of previous morbidity surveys (e.g. Canada Health Survey, 1981; Branch & Jette, 1981; Jette and Branch, 1981; Nathanson, 1977), the variables age and sex were not significantly correlated with objective or subjective health status. That is, when considered independently or jointly by age/sex categories, there were no significant relationships between the age and sex of the respondents and the number of diagnoses, the number of active problems listed, self-ratings of health, or total SIP scores. Table V.3 summarizes the interrelationships of these variables.

T-tests for independent groups revealed no significant differences between males and females in mean score across SIP categories or dimensions. The same finding was evident when individuals in the age range 65 to 74 years were compared with those 75 years of age and older, and when combined age/sex categories defined subgroups for comparison.

These results are due perhaps to the general health status of the sample, insofar as all participants in the study required medical, nursing, rehabilitation, and/or social services. The sample was very homogeneous in this regard, which in turn reflected the criteria for admission to the Day Hospital.

TABLE V.3
CORRELATIONS BETWEEN AGE, SEX,
NO. DIAGNOSES, NO. ACTIVE PROBLEMS,
SELF-RATED HEALTH, AND OVERALL SIP

	AGE	SEX	NO. DIAG.	NO. PROB.	SRH
SEX	-0.09				
NO.DIAG.	-0.01	-0.01			
NO.PROB.	-0.02	0.10	-0.01		
SRH	-0.18	0.09	-0.06	0.16	
OVERALL SIP	-0.10	0.11	-0.07	0.63**	0.40*

* $p < .05$
** $p < .01$

The small size of the sample (N=30), consisting of volunteers unrepresentative of the elderly population of the city of Edmonton, also limits the generalizability of the findings when considering age/sex differentials in morbidity. However, it is suggested that age and sex may not be important predictors of health status in select patient groups such as the one used in this study.

D. Outcome of the Day Hospital Program as Measured by the SIP

The outcome of the Day Hospital program was measured in terms of changes in the perceived health status of the participants in the study, as indicated by SIP scores obtained on two occasions. Among the thirty members of the sample, only four people did not improve in perceived health status between the first and second administrations of the SIP. Six respondents demonstrated no change, as evidenced by

minimal (ie. less than 2.00 point) fluctuations in overall SIP scores during the same time period.

Paired t-tests for category, dimension, and overall SIP scores were conducted to determine the significance of the observed changes. With reference to Table V.4, significant differences existed for scores in the categories Emotional Behaviour, Alertness Behaviour, Body Care and Movement, and Recreation and Pastimes, as well as in the psychosocial and physical dimensions and the overall SIP. All significant changes were in the direction of improved health status.

Given that the sample initially was most dysfunctional in terms of Recreation and Pastimes and Alertness Behaviour, it may be said that the Day Hospital program was somewhat successful in reducing the impact of health-related problems in these areas. Further, if one considers the level of significance obtained for the changes in category scores, it is tempting to speculate that the most dramatic improvements in perceived health status were in the psychosocial domain. Indeed, as the participants in the study had predominantly chronic physical illnesses, great changes in physical health status were not anticipated. However, with interventions directed toward minimizing the impact of disease upon functional abilities, changes in mental outlook toward one's health seemed more likely to occur and to be expressed as changes in perceived health status.

TABLE V.4
PAIRED T-TESTS FOR SIP CATEGORY, DIMENSION, AND OVERALL SCORES, T1 AND T2

TIME	VAR.	NO. CASES	MEAN	(SD)	t	PROB.	TEST RETEST CORREL.
T1	SR	30	21.04	(21.28)	0.80	N.S.	0.81
T2			19.12	(20.90)			
T1	EB	30	25.86	(26.44)	2.80	p < .01	0.85
T2			18.77	(22.71)			
T1	BCM	30	17.92	(15.47)	2.16	p < .05	0.84
T2			14.57	(13.25)			
T1	HM	30	31.07	(24.55)	0.21	N.S.	0.72
T2			30.37	(23.26)			
T1	M	30	22.39	(14.47)	1.78	N.S.	0.46
T2			17.67	(13.52)			
T1	SI	30	16.44	(10.63)	1.65	N.S.	0.60
T2			13.64	(10.28)			
T1	A	30	26.34	(16.70)	1.47	N.S.	0.85
T2			23.90	(15.93)			
T1	AB	30	26.92	(30.56)	2.94	p < .01	0.68
T2			18.12	(27.03)			
T1	C	30	13.19	(14.50)	2.94	p < .01	0.75
T2			7.94	(8.9)			
T1	W	30	0.50	(2.73)	0.00	N.S.	1.00
T2			0.50	(2.73)			
T1	RP	30	33.12	(25.84)	2.04	p < .05	0.61
T2			25.27	(20.41)			
T1	E	30	6.75	(5.11)	0.95	N.S.	0.76
T2			6.13	(5.14)			
T1	Physical	30	20.81	(12.70)	2.66	p < .05	0.84
T2	Dimension		17.40	(12.11)			
T1	Psychosocial	30	19.84	(15.71)	3.01	p < .01	0.78
T2	Dimension		14.45	(12.82)			
T1	Overall	30	19.70	(11.03)	3.08	p < .01	0.80
T2	SIP Score		16.00	(9.48)			

E. Reliability of SIP Scores

Test-retest reliability estimates for overall, dimension, and category scores were based on the correlation between results of of the first and second administrations of the SIP. With reference to Table V.4, test-retest reliability for overall and dimension scores were within acceptable limits and reflected the tendency for SIP scores to be stable over the one month period between administrations. Notably, test-retest score correlations ranged between 0.46 (Mobility) and 1.00 (Work) for the categories; the reliability of each category is discussed in the section "SIP Category Analysis".

Reliability of the SIP in terms of internal consistency was assessed by Cronbach's (1951) coefficient alpha. High values of coefficient alpha indicate an instrument measures traits that are the same or that are highly correlated (Allen & Yen, 1979).

In terms of the overall SIP, alpha coefficients were based on the relationship among categories. With reference to Table V.5, the internal consistency of the SIP was 0.93 overall, and thus it appears very homogeneous in the constructs it measures. The psychosocial dimension possessed a high degree of internal consistency (α=.90), but the physical dimension (α =.84) was less homogeneous as categories M and A did not correlate significantly within it.

TABLE V.5
INTERNAL CONSISTENCY OF DIMENSION AND OVERALL SCORES OF THE SIP AS INDICATED BY CRONBACH'S (1951) ALPHA

	ALPHA
Physical Dimension	0.84
Psychosocial Dimension	0.90
Overall SIP	0.93

Table V.6 gives the item/category correlations and coefficient alpha for each category of the SIP. Internal consistency was quite variable across categories, with a range from 0.10 (E) to 0.86 (AB). The homogeneity of each category was based on scaled item responses and is discussed in the section "SIP Category Analysis".

F. Validity of SIP Scores

An indicator or index is a valid measure of health status only to the extent that it expresses or quantifies that construct (Kaplan et al., 1976). Although establishing construct validity is an unending process (Allen & Yen, 1979), it may be estimated by observing the relationships between an instrument and other indicators of the same characteristics.

The validity of the SIP was assessed by observing its relationship to physician diagnoses, active problems on problem lists of medical records, and self-ratings of

TABLE V.6
ITEM/CATEGORY CORRELATIONS AND INTERNAL CONSISTENCY OF SIP CATEGORIES

CATEGORY	NO. ITEMS	ITEM / CORRELATIONS					ALPHA
		(0.0 to 0.20)	(0.21 to 0.40)	(0.41 to 0.60)	(0.61 to 0.80)	(0.81 to 1.00)	
BCM	23	n = 6	n = 6	n = 10	n = 1	n = 0	0.82
M	10	n = 6	n = 3	n = 1	n = 0	n = 0	0.40
A	12	n = 5	n = 2	n = 2	n = 3	n = 0	0.61
EB	9	n = 0	n = 1	n = 6	n = 2	n = 0	0.82
SI	20	n = 8	n = 7	n = 4	n = 1	n = 0	0.58
AB	10	n = 0	n = 0	n = 8	n = 1	n = 1	0.86
C	8	n = 3	n = 3	n = 3	n = 0	n = 0	0.57
SR	7	n = 2	n = 2	n = 2	n = 1	n = 0	0.68
HM	10	n = 3	n = 2	n = 3	n = 2	n = 0	0.72
W	8	n = 7	n = 0	n = 0	n = 0	n = 2	0.55
RP	8	n = 1	n = 3	n = 4	n = 0	n = 0	0.68
E	9	n = 8	n = 1	n = 0	n = 0	n = 0	0.10

health. Table V.3 summarizes the interrelationships of these variables, each of which is discussed separately as follows.

Relationship with Physician Diagnoses

The number of diagnoses made by the physician for each participant in the study was not significantly correlated with overall SIP scores. However, evidence was found for the ability of the SIP to discriminate between diagnostic subgroups within the sample. That is, the descriptive validity of the SIP, defined as "the extent to which dimension and category scores and item-checking patterns provide a useful and meaningful qualitative description of different samples and types of subjects" (Bergner et al., 1981, p. 799), was substantiated insofar as depressed members of the sample were concerned. With reference to Table V.7 and Figure V.1, those members of the sample (n=15) with the diagnosis of depression were significantly different from the rest of the sample in terms of mean scores in the psychosocial dimension and in three of the component categories of this domain, Emotional Behaviour, Social Interaction, and Alertness Behaviour. Further, items in categories EB, SI, and AB checked by most of the "depressed" respondents served to clarify the nature of their psychosocial dysfunction. Table V.8 lists these items, while a list of the items frequently checked by this group across all SIP categories is in Appendix G.

TABLE V.7
COMPARISON OF DEPRESSED PATIENTS* WITH THE REST
OF THE SAMPLE** ACROSS CATEGORY, DIMENSION, AND OVERALL SIP SCORES

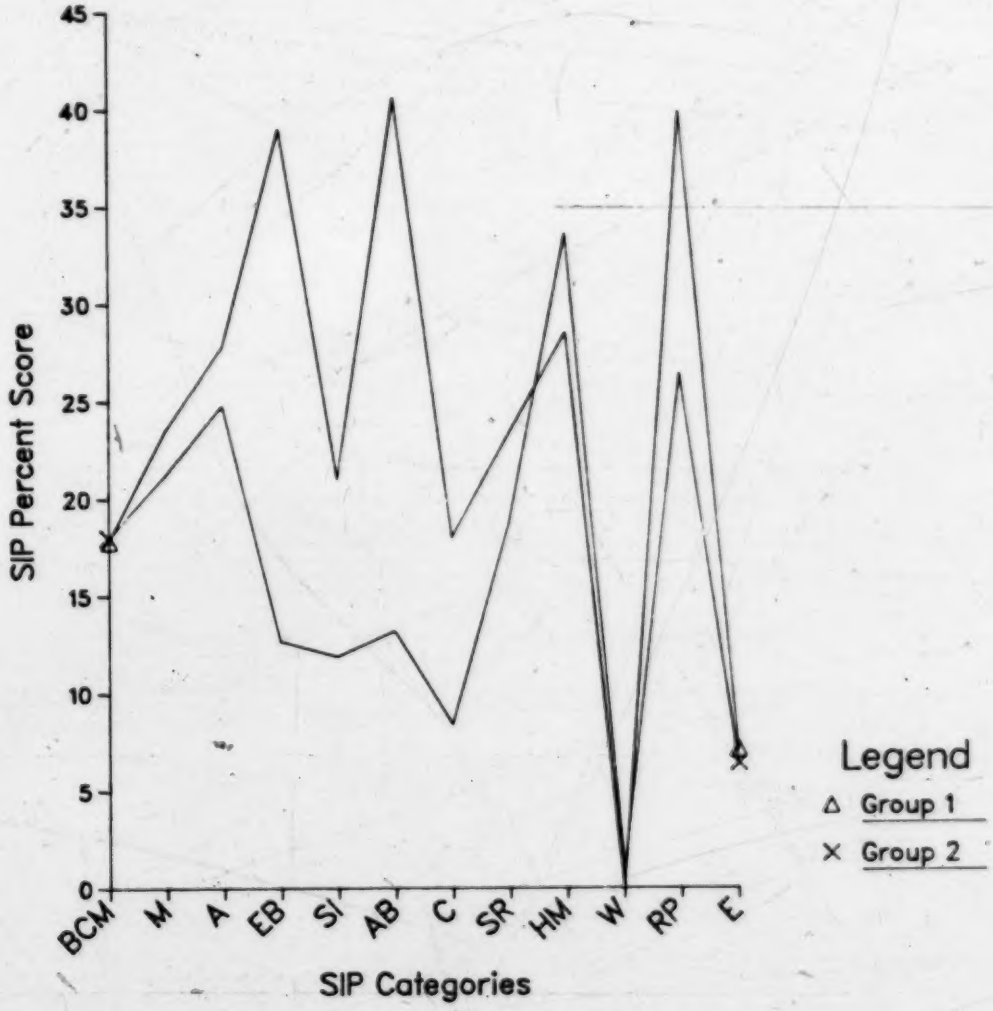
CATEGORY / DIMENSION	GROUP	MEAN	(S.D.)	t	PROB.
SR	I	23.3	(25.6)	0.59	N.S.
	II	18.7	(16.5)		
EB	I	38.9	(30.5)	3.10	p < .01
	II	12.7	(12.2)		
BCM	I	17.8	(13.2)	-0.03	N.S.
	II	18.0	(17.9)		
HM	I	28.5	(20.4)	-0.57	N.S.
	II	33.6	(28.6)		
M	I	23.5	(13.0)	0.43	N.S.
	II	21.2	(16.2)		
SI	I	21.0	(12.2)	2.55	p < .05
	II	11.9	(6.5)		
A	I	27.8	(16.8)	0.49	N.S.
	II	24.8	(17.1)		
AB	I	40.6	(33.7)	2.71	p < .05
	II	13.2	(19.9)		
C	I	18.0	(17.8)	1.91	N.S.
	II	8.4	(8.2)		
RP	I	39.9	(25.5)	1.46	N.S.
	II	26.4	(25.2)		
E	I	7.1	(3.8)	0.34	N.S.
	II	6.4	(6.2)		
PHYS	I	21.3	(11.7)	0.23	N.S.
	II	20.3	(14.1)		
PSYCH	I	28.0	(17.9)	3.31	p < .01
	II	11.6	(6.8)		
OVERALL SIP SCORE	I	23.0	(12.1)	1.72	N.S.
	II	16.4	(9.0)		

* GROUP I (N = 15)
** GROUP II (N = 15)

TABLE V.8
HIGHLY PREVALENT BEHAVIOURAL DYSFUNCTIONS OF DEPRESSED PATIENTS (n=15) AS MEASURED IN CATEGORIES EB, SI, AND AB OF THE SIP

CATEGORY	ITEM	NO. (%) CHECKING ITEM
EB	"I act nervous or restless."	12 (80%)
	"I talk about the future in a hopeless way."	8 (53.3%)
SI	"I am going out less to visit people."	12 (80%)
	"I am doing fewer social activities with groups of people."	10 (66.7%)
	"I often express concern over what might be happening to my health."	9 (60%)
AB	"I have difficulty reasoning and solving problems, for example, making plans, making decisions, learning new things."	9 (60%)
	"I make more mistakes than usual."	8 (53.3%)
	"I have difficulty doing activities involving concentration and thinking."	8 (53.3%)

SIP SCORES OF DEPRESSED PATIENTS COMPARED TO REST OF SAMPLE



GROUP 1: DEPRESSED PATIENTS (N=15)

GROUP 2: ALL OTHER MEMBERS OF THE SAMPLE (N=15)

Relationship with Active Problem List

The construct validity of the SIP as a health status measure was further substantiated by its relationship to the number of active problems listed in problem oriented medical records. Pearson correlation coefficients between the number of active problems and the physical dimension ($r=0.63$), the psychosocial dimension ($r=0.49$), and overall SIP scores ($r=0.63$) were all significant at the $p<.01$ level. In addition, stepwise multiple regression analyses indicated that among the independent variables of age, sex, number of diagnoses, and self-ratings of health, the number of active problems listed was the only significant predictor of overall SIP scores, accounting for 29.86% of the variance.

If one considers the number of problems listed and the number of items checked on the SIP as both on a continuum reflecting mild to severe impairment of health, then a positive relationship between them may be predicted. As a problem must be validated by the patient in order to be included in the problem list of his medical record, it is forseeable that the number of problems thus identified at any point in time will positively correlate with the number of behavioural dysfunctions (items) checked on the SIP (and hence with total SIP score).

Relationship with Self-Rated Health

Consistent with their mutual status as subjective measures of health, self-ratings of health correlated

significantly with overall SIP scores ($r=0.40$, $p<.05$). Thus, the construct validity of the SIP as a measure of perceived health status was supported.

It was interesting to note the significant correlation ($r=0.38$, $p<.05$) between self-ratings of health and the psychosocial dimension of the SIP. Categories EB ($r=0.54$, $p<.01$) and SI ($r=.34$, $p<.05$) of this dimension, as well as category BCM ($r=0.39$, $p<.05$) of the physical dimension, also were significantly related to self-ratings of health. Further, category EB was the only significant predictor of self-rated health when stepwise multiple regression techniques were introduced, independently accounting for 29% of the variance.

Exploration of the relationships between self-ratings of health and the Emotional Behaviour category of the SIP led to the observation that individuals who rated their health as "poor" (n=6) checked the highest proportion of items in this category. That is, those people who said their health was "poor" also tended to indicate worse health in terms of Emotional Behaviour than individuals who said their health was "fair" or "good". As evidence of the descriptive capability of the SIP, Table V.9 lists the items in category EB checked frequently by these six people, and Appendix H contains the behavioural dysfunctions identified by all members of this subgroup as measured across SIP categories.

Self-ratings of health of individuals with the diagnosis of depression (n=15) were then examined, as it was

TABLE V.9
HIGHLY PREVALENT BEHAVIOURAL DYSFUNCTIONS OF RESPONDENTS
RATING THEIR HEALTH AS POOR (n = 6), AS MEASURED BY CATEGORY EB OF THE SIP

ITEM IN CATEGORY EB OF THE SIP	NO. (%) CHECKING ITEM
1. "I say how bad or useless I am, for example, that I am a burden on others".	6 (100.0%)
3. "I often moan and groan in pain or discomfort".	5 (83.3%)
5. "I act nervous or restless".	4 (66.7%)
6. "I keep rubbing or holding areas of my body that hurt or are uncomfortable"	6 (100.0%)
8. "I talk about the future in a hopeless way".	4 (66.7%)

known they had significantly higher mean scores than the rest of the sample (n=15) in categories EB, SI, and AB of the SIP. Figure V.2 and Table V.10 illustrate the distribution of self-rated health for these two subgroups. The modal self-rating of health for the "depressed" group was "fair", while the rest of the sample most frequently rated their health as "good".

Analysis of variance on the overall SIP scores of three subgroups of the sample, defined on the basis of self-rated health, revealed significant differences between them ($F=9.19$, $p<.001$). When the sample was partitioned into two groups, defined on the basis of presence or absence of diagnosed depression, and mean overall SIP scores were compared across categories of self-rated health, the "depressed" group consistently had higher SIP scores. Table V.11 summarizes this data.

Taken together, these findings provide considerable evidence for a positive relationship between relatively negative self-ratings of health, higher SIP scores in the psychosocial domain, and depression. In summary, there was a significant, positive correlation between overall SIP scores and self-ratings of health, between psychosocial dimension scores of the SIP and self-ratings of health, and between categories EB, SI, and BCM of the SIP and self-ratings of health. Category EB was the only significant predictor of self-rated health according to the multiple regression analyses conducted, and items in this category are checked

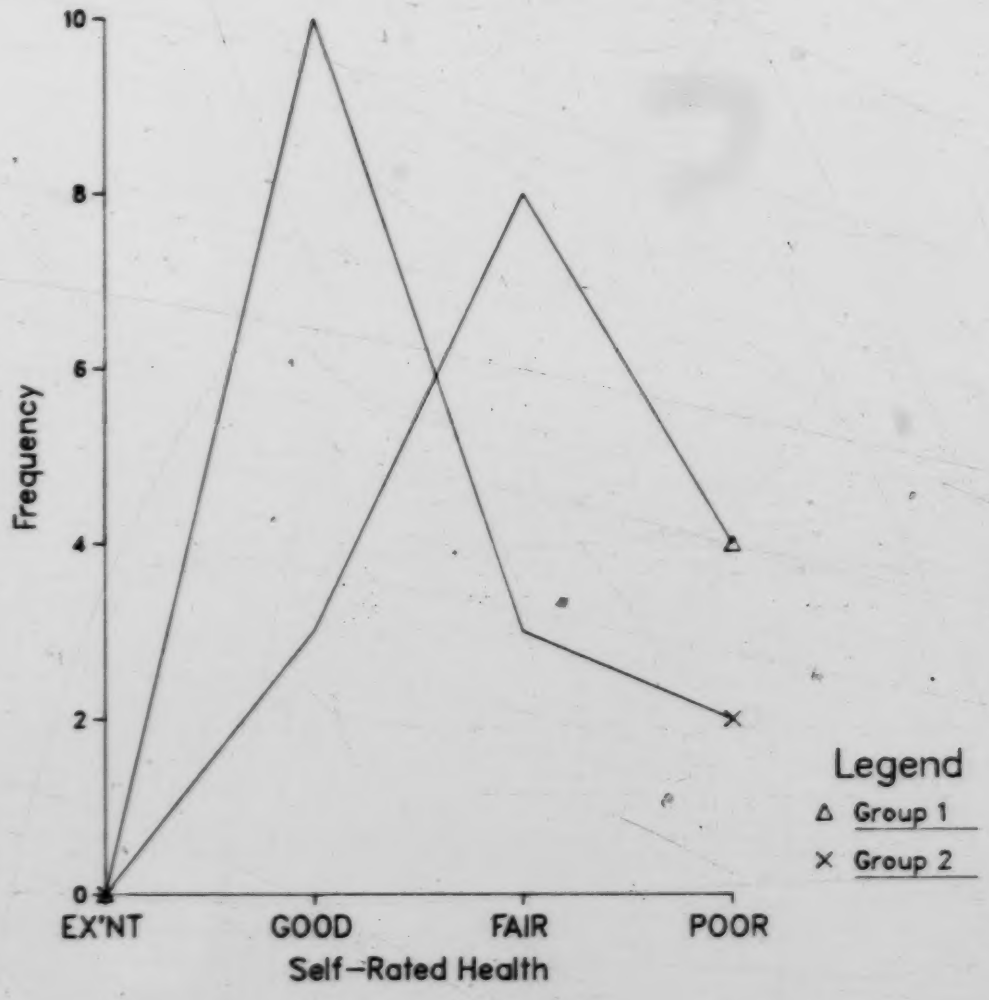
TABLE V.10

OVERALL SIP SCORES BY SELF-RATINGS OF HEALTH FOR TOTAL SAMPLE AND SUBGROUPS DEFINED ON THE BASIS OF DEPRESSION

SELF-RATING OF HEALTH	OVERALL SIP SCORE					
	TOTAL SAMPLE (N=30)		GROUP I* (N=15)		GROUP II** (N=15)	
	MEAN	(S.D.)	MEAN	(S.D.)	MEAN	(S.D.)
EXCELLENT	0.0	( )	0.0	( )	0.0	( )
GOOD	18.2	(9.7)	28.0	(8.9)	15.3	(8.2)
FAIR	14.2	(8.3)	15.4	(9.1)	10.8	(5.3)
POOR	33.0	(7.4)	34.5	(9.1)	30.1	(0.1)

* GROUP I: Individuals with diagnosis of depression (n = 15).
** GROUP II: Individuals without diagnosis of depression (n = 15).

SELF–RATED HEALTH OF DEPRESSED PATIENTS *



*GROUP 1: DEPRESSED PATIENTS (N=15)

GROUP 2: ALL OTHER MEMBERS OF THE SAMPLE (N=15)

TABLE V.11
SELF-RATINGS OF HEALTH
BY DEPRESSED PATIENTS* (n = 15)
COMPARED WITH THE REST OF THE SAMPLE** (n = 15)

	SELF-RATING OF HEALTH				
	EXC.	GOOD	FAIR	POOR	TOTAL
GROUP I*	0	3	8	4	15
GROUP II**	0	10	3	2	15
TOTAL	0	13	11	6	N=30

most frequently by individuals who rate their health as "poor". Depressed members of the sample had significantly higher mean scores in categories EB, SI, and AB of the psychosocial dimension of the SIP, and they usually rated their health as "fair". Further, the "depressed" subgroup of the sample had higher mean overall SIP scores across all categories of self-rated health. That is, regardless of how they rated their health, these individuals perceived their health to be worse than members of the sample not diagnosed as being depressed.

Descriptive Validity

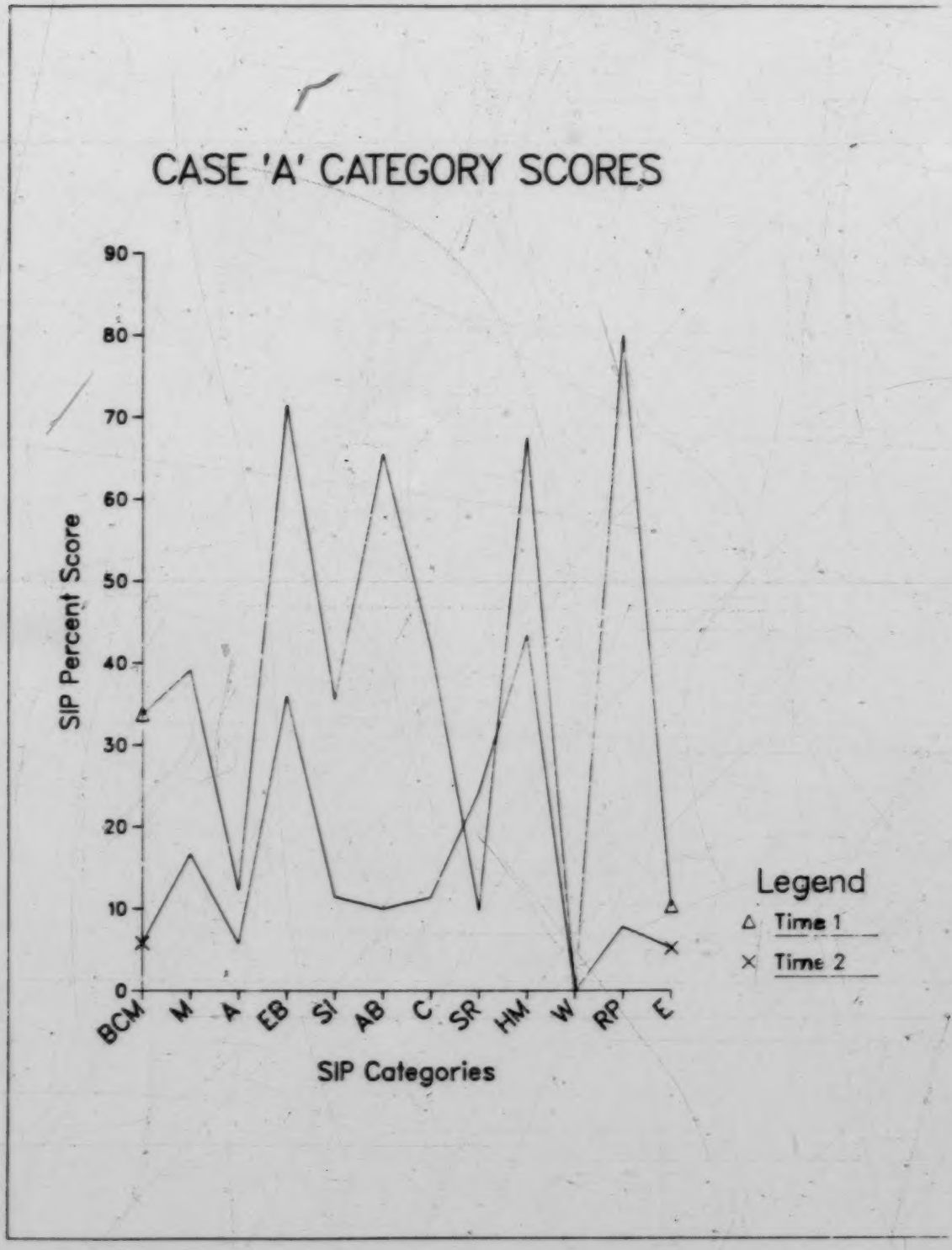
The ability of the SIP to discriminate depressed patients from the rest of the sample, as well as its descriptive validity in terms of the specific behavioural dysfunctions of this group, has already been noted. As evidence of the validity of the SIP as a descriptor of

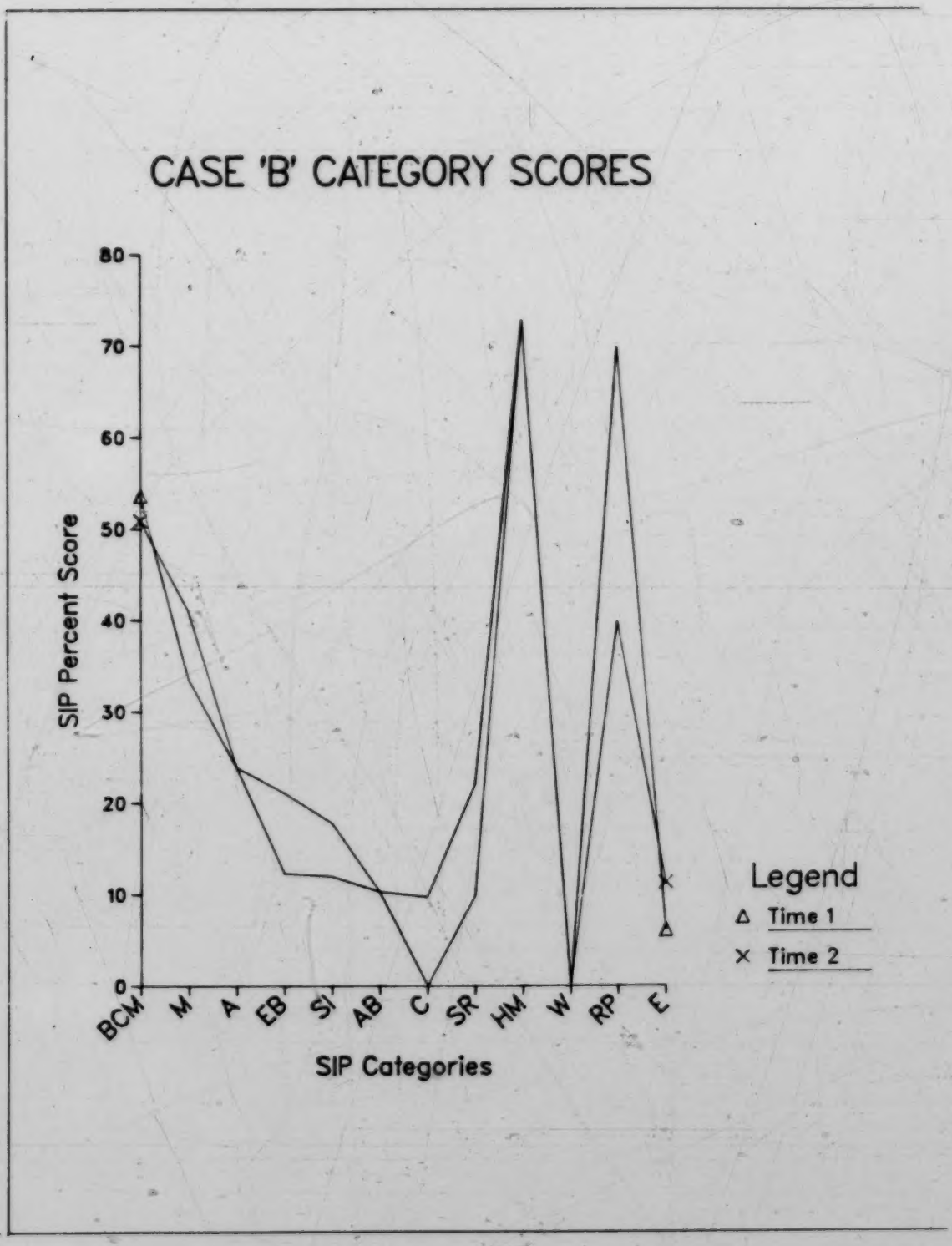
individual patients, the following case examples are presented.

Case 'A' (Figure V.3) is a 75 year old woman with a history of recent bowel surgery, depression and anxiety, osteoarthritis, and seizures of unknown origin. Low affect, joint pain, decreased mobility, and decreased ADL (activities of daily living) skills were among the active problems listed at the time of the first interview.

Initial administration of the Sickness Impact Profile revealed severe dysfunction in the psychosocial domain, especially as measured by categories EB and AB. Problems were also perceived in terms of home management and recreation and pastimes. By the time of the second interview, however, considerable improvement across all areas was evident. This finding paralleled clinical observations which emphasized improvements in affect and ADL skills. It was apparent in this case that improvements in the psychosocial domain had a generalized influence on perceived health status.

Case 'B' (Figure V.4) is a 66 year old female with rheumatoid arthritis. Active problems listed were decreased strength and range of motion in all muscle groups and joints, decreased functional abilities in activities of daily living, joint pain, and edema of legs and ankles. SIP scores obtained one month apart revealed persistent


CASE 'A' CATEGORY SCORES
SIP Percent Score
90
80
70
60
50
40
30
20
10
0
BCM
M
A
EB
SI
AB
C
SR
HM
W
RP
E
SIP Categories
Legend
Time 1
Time 2


CASE 'B' CATEGORY SCORES
SIP Percent Score
80
70
60
50
40
30
20
10
0
BCM
M
A
EB
SI
AB
C
SR
HM
W
RP
E
SIP Categories
Legend
Time 1
Time 2

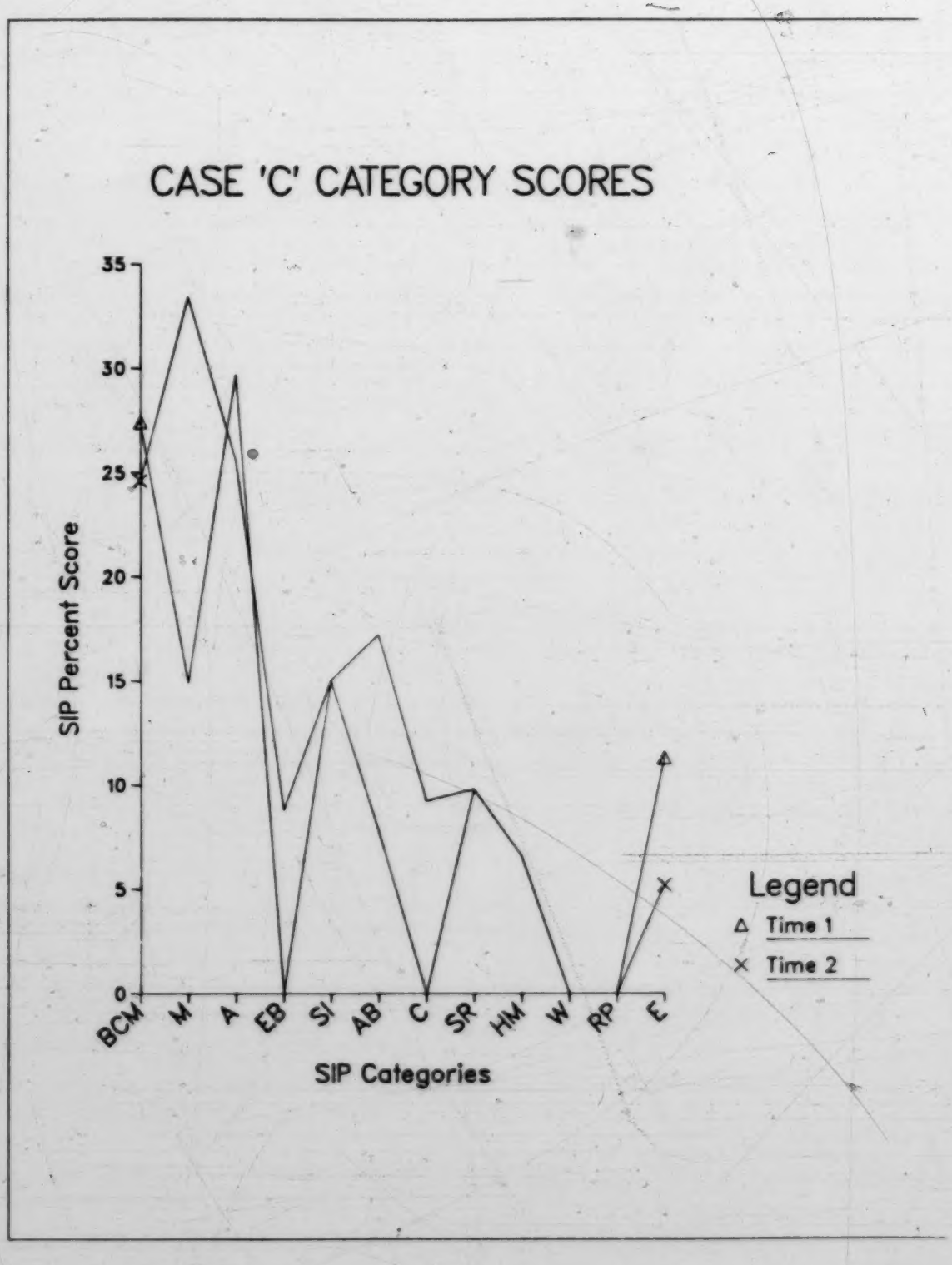
elevation of scores in categories Body Care and Movement, Home Management, and Recreation and Pastimes. Her scores reflected the impact of chronic disease upon her physical health status. Consistent with the findings of Deyo et al. (1983), scores in categories of the psychosocial dimension were considerably lower than those indicating physical dysfunction.

Case 'C' (Figure V.5) was a 91 year old male with osteoarthritis, ischaemic heart disease, cataracts, bronchiectasis, and peripheral neuropathy. Active problems identified were decreased endurance, instability of the right knee, poor balance, poor vision, and impaired sensation in his feet.

The SIP's in this case confirmed clinical reports that dysfunction was essentially limited to the physical domain. Highest category scores were in Body Care and Movement, and Ambulation, respectively. The elevated score in category M upon the second administration of the SIP reflected the generalized impact of his difficulty with walking upon his mobility in the community.

G. SIP Category Analysis

Each of the twelve categories of the SIP is intended to measure a distinct area of living that may be subject to the impact of sickness. To the extent that these categories do not correlate with each other, it may be said that they


CASE 'C' CATEGORY SCORES
SIP Percent Score
35
30
25
20
15
10
5
0
BCM
M
A
EB
SI
AB
C
SR
HM
W
RP
E
SIP Categories
Legend
Time 1
Time 2

contribute unique information about the perceived health status of an individual or group. However, it is also recognized that the more dysfunctional the sample, the more categories are responded to, and the higher the intercorrelation of category scores and the correlation of each category with overall SIP score (Bergner et al., 1981).

In the sections that follow, each SIP category is described in terms of its interrelationships with overall SIP score, dimension scores, and the other categories. The test-retest reliability, internal consistency, and validity of each category are also examined. Three sections follow consisting of the independent categories, the categories of the physical dimension, and the categories of the psychosocial dimension. Correlations between all categories are shown in Table V.12.

Independent Categories

With reference to Table V.13, Categories SR, HM, W, RP, and E were, with one exception, not significantly correlated with each other. This indicated that each contributed unique information regarding the perceived health status of the sample. Stepwise multiple regression analyses confirmed that each of these categories independently accounted for some portion of the variance in overall SIP scores.

Sleep and Rest

The Sleep and Rest (SR) category was positively correlated with overall SIP scores ($r=0.62$, $p<.001$). It also

TABLE V.12
INTERRELATIONSHIPS OF SIP CATEGORIES

CATEGORY	BCM	M	A	EB	SI	AB	C	SR	HM	W	RP
BCM											
M	0.40*										
A	0.49**	0.32									
EB	0.45*	0.35	0.39*								
SI	0.42	0.47	0.11	0.76***							
AB	0.10	0.37	0.11	0.70***	0.66***						
C	0.44	0.62	0.32	0.42*	0.50**	0.36*					
SR	0.33	0.51**	0.32	0.54**	0.60**	0.40*	0.25				
HM	0.72***	0.46**	0.30	0.33	0.39*	0.16	0.47	0.28			
W	0.12	-0.10	0.04	-0.18	-0.03	-0.12	-0.17	-0.10	-0.19		
RP	0.42*	0.15	0.09	0.23	0.43*	0.13	0.58**	0.11	0.42*	-0.24	
E	0.16	0.12	0.14	0.20	0.30	0.19	0.02	0.21	0.15	0.17	0.06

* p < .05
** p < .01
*** p < .001

TABLE V.13
INTERRELATIONSHIPS OF SIP INDEPENDENT CATEGORIES

	SR	HM	W	RP	E
SR					
HM	N.S.				
W	N.S.	N.S.			
RP	N.S.	0.42*	N.S.		
E	N.S.	N.S.	N.S.	N.S.	
OVERALL SIP	0.62***	0.69***	N.S.	0.49**	N.S.

* p < .05
** p < .01
*** p < .001

had significant relationships with M (r=0.51, p<.01), EB (r=0.54, p<.01), SI (r=0.60, p<.01), and AB (r=0.40, p<.05). Thus, apparently disturbances in sleep and rest behaviour especially accompany dysfunction in the psychosocial domain. Test-retest reliability was high (r=0.81), while internal consistency was somewhat lower (α=0.68). Elevated SIP scores in this category were generally validated by clinical reports of problematic sleep patterns.

Home Management

The Home Management (HM) category was positively correlated with overall SIP score (r=0.69, p<.001). It also had significant relationships with two categories of the physical dimension, BCM (r=0.72, p<.001) and M (r=0.46, p<.01), and with category SI of the psychosocial dimension (r=0.39, p<.05). Test-retest reliability as well as internal consistency were estimated as 0.72. Item-checking patterns

in this category appeared to be influenced by the availability of social support systems.

Work

The category Work (W) was generally not applicable to this sample, as only one participant in the study was not retired. Thus, all but this person obtained a score of 0.0 in the category.

W was not significantly correlated with overall SIP scores ($r=-0.08$) or with any single category of the instrument. Test-retest reliability was 1.00, indicating that the sole member of the sample who worked continued to do so and that he checked the same items in this category on both administrations of the SIP. Coefficient alpha was 0.55; seven of the nine items were weakly correlated with category scores and the remaining two had item/category correlations in the range of 0.81 to 1.00. These correlation coefficents were not considered valid indicators of the internal consistency of this category, as they were based on the item-checking pattern of only one person. Nevertheless, as a general statement it may be said that this category had minimal validity when used with this elderly group.

Recreation and Pastimes

The category Recreation and Pastimes (RP) was positively correlated with overall SIP scores ($r=0.49$, $p<.01$). It also had significant relationships with BCM ($r=0.42$, $p<.05$), SI ($r=0.43$, $p<.05$), C ($r=0.58$, $p<.01$), and

HM ($r=0.42$, $p<.05$). These interrelationships were seen as indicators of the general level of dysfunction in the sample. Further, the fact that the highest mean category score for the sample was in RP may be due to the retirement status of the respondents, coupled with social isolation arising from impaired health.

Test-retest reliability was 0.61, indicating that scores were somewhat stable. Internal consistency was estimated as 0.68, suggesting that RP measures more than one construct.

Eating

The category Eating (E) was not significantly correlated with overall SIP score or any single category of the instrument. Test-retest reliability was moderate ($r=0.76$), but internal consistency was extremely low ($\alpha=0.10$). With reference to Appendix F, five of its nine items were not checked by any member of the sample and these items described the most severe dysfunctions in eating behaviour.

Notably, during the development of the SIP, Bergner et al. (1981) found that items concerned with eating and taking nutrition did not make a statistically significant contribution to the SIP. However, as "consideration of the application of the SIP in clinical and program evaluation settings indicated that this category made a substantive contribution to the descriptive capacity of the instrument" (p. 803), this category was retained in the final revision

of the SIP. It would appear that similar considerations have surfaced in this study.

Physical Dimension

The categories BCM, M and A are aggregated into the "physical dimension" of the SIP. With reference to Table V.14, the interrelationships of these categories with physical dimension and overall SIP scores are summarized. Stepwise multiple regression analyses indicated that BCM was the strongest predictor of physical dimension scores, independently accounting for 88% of the variance. It was followed by Ambulation (7%) and Mobility (5%), respectively.

Test-retest reliability of the physical dimension was moderate (0.78), while internal consistency was high (r=0.90). Notably, physical dimension scores were significantly correlated with psychosocial dimension scores ($r=0.46$, $p<.05$), suggesting that worse physical function is accompanied by worse psychosocial function at a given point in time.

Body Care and Movement

The Body Care and Movement (BCM) category was strongly correlated with physical dimension scores ($r=0.93$, $p<.001$), and moderately related to overall SIP scores ($r=0.77$, $p<.001$). As noted, BCM was the strongest predictor of physical dimension scores, accounting for 88% of the variance. It had significant relationships with the two other categories of this dimension, as well as with category

TABLE V.14
CORRELATION MATRIX OF SIP PHYSICAL DIMENSION

	BCM	M	A	PHYS
BCM				
M	0.40*			
A	0.49**	N.S.		
PHYS DIMEN.	0.93***	0.60***	0.72***	
OVERALL SIP	0.77***	0.64***	0.52**	0.84***

* $p < .05$
** $p < .01$
*** $p < .001$

EB of the psychosocial dimension (r=0.45, p<.05) and the independent categories HM (r=0.72, p<.001) and RP (r=0.42, p<.05).

Test-retest reliability was high (r=0.84), and BCM was generally homogeneous as indicated by a coefficient alpha of 0.82.

Mobility

The category Mobility (M) was moderately correlated with the physical dimension (r=0.60, p<.001) and overall SIP scores (r=0.64, p<.001). It was a weak predictor of physical dimension scores, however, and was not significantly correlated with category A of this dimension. Mobility did correlate significantly, nevertheless, with categories SR (r=0.51, p<.01) and HM (r=0.46, p<.01). Inspection of items contained within these categories suggested a common underlying factor based on patterns of rest and activity.

Test-retest reliability of M was low (r=0.46), as was its coefficient alpha (α=0.40). Indeed, according to the comments of respondents, item-checking patterns in this category were influenced by physical ability to walk, sleep and rest routines, opportunities for socialization, vision, and the availablity of transportation.

Ambulation

The Ambulation (A) category was positively correlated with physical dimension (r=0.72, p<.001) and overall SIP scores (r=0.52, p<.01). It had significant relationships with BCM (r=0.49, p<.01) within the dimension, and with EB (r=0.39, p<.05) of the psychosocial dimension.

Test-retest reliability was high (r=0.85), while it was somewhat homogeneous in terms of constructs measured (α=0.61). Elevated scores in this category were consistently validated by clinical reports of patient difficulty with ambulation.

Psychosocial Dimension

The categories EB, SI, AB, and C are aggregated into the "psychosocial dimension" of the SIP. With reference to Table V.15, the interrelationships of these categories with psychosocial dimension and overall SIP scores are summarized.

Stepwise multiple regression analyses indicated the psychosocial dimension was the strongest predictor of overall SIP scores, independently accounting for 71% of the

TABLE V.15
CORRELATION MATRIX OF SIP PSYCHOSOCIAL DIMENSION

	EB	SI	AB	C	PSYCH
EB					
SI	0.78***				
AB	0.70***	0.66***			
C	0.42*	0.50*	0.36*		
PSYCHO-SOC. DIM.	0.89***	0.88***	0.88***	0.60***	
OVERALL SIP	0.79***	0.80***	0.61***	0.68***	0.85***

* $p < .05$
** $p < .01$
*** $p < .001$

variance. Further, category EB was the strongest predictor of psychosocial dimension scores, independently accounting for 77% of the variance. It was followed by categories AB (14%), SI (6%), and C (3%), respectively.

Test-retest reliability of the psychosocial dimension was moderate (r=0.78), while its internal consistency was high (α=0.90). The clinical validity of this dimension was especially substantiated in terms of depressed patients, and it was seen to have interesting significant relationships with self-ratings of health.

Emotional Behaviour

The Emotional Behaviour (EB) category was positively correlated with psychosocial dimension (r=0.89, p<.001) and overall SIP scores (r=0.79, p<.001). It had significant relationships with all other categories in the psychosocial

dimension, as well as with A (r=0.39, p<.05) of the physical dimension, and SR (r=0.54, p<.01). As noted, EB was the strongest predictor of psychosocial dimension scores, and it was the only significant predictor of self-ratings of health, accounting for 29.8% of the variance.

Test-retest reliability was high (r=0.85), and it was generally homogeneous in the constructs measured (α =0.82). Clinical validity of this category as a descriptor of depressed patients and of those who rated their health as "poor" was also apparent.

Social Interaction

The Social Interaction (SI) category was positively correlated with psychosocial dimension (r=0.88, p<.001) and overall SIP scores (r=0.80, p<.001). It had significant relationships with all other categories in the psychosocial dimension as well as with HM (r=0.39, p<.05) and RP (r=0.43, p<.05).

Test-retest reliability was moderate (r=0.60), while internal consistency was lacking (α=0.58).

Alertness Behaviour

The Alertness Behaviour (AB) category was positively correlated with psychosocial dimension (r=0.88, p<.001) and overall SIP scores (r=0.61, p<.001). It had significant relationships with all other categories in this dimension, but did not overlap to any significant degree with any other category of the instrument. The mean score in this category

was the second highest for this sample.

Test-retest reliability was moderate (r=0.68), while it was the most internally consistent category of the SIP (r=0.86).

Communication

The Communication (C) category had positive relationships with psychosocial dimension (r=0.60, p<.001) and overall SIP scores (r=0.68, p<.001). It was significantly related to all other categories of the psychosocial dimension as well as with the independent category RP (r=0.58, p<.01).

Test-retest reliablity was moderate (r=0.75), while it lacked internal consistency (α=0.68). Items in this category were not frequently checked by members of the sample, with the exception of the item describing difficulty with writing or typing.

H. Participant Evaluation of the SIP

Table V.16 summarizes responses to questions posed during the participant interviews. In general, respondents considered the SIP to give an accurate picture of how they saw their health.

Only one person reported difficulty in understanding the instructions for answering the SIP; this was remedied by repeating the directions as needed. However, among those who reported no difficulty in understanding the instructions were individuals who required assistance to read SIP

TABLE V.16
EVALUATION OF THE SIP BY RESPONDENTS (N = 30)

CRITERIA (PARAPHRASED)	YES (%)	NO (%)
1. SIP gives accurate picture of my health.	27(90.0%)	3(10.0%)
2. Had difficulty in understanding instructions for answering SIP	1(3.3%)	29(96.7%)
3. Willing to share SIP information with health professionals of Day Hospital	29(96.7%)	1(3.3%)

statements because of poor vision or illiteracy. For those with hearing problems, it was necessary to read each item in a loud voice.

Respondents were almost unanimous in their willingness to share their SIP data with the health care professionals of the Day Hospital. The sole person who refused permission stated, "They should know this information already". Thus, generally speaking, the SIP may be an acceptable tool for increasing the participation of elderly patients in health care processes.

Nevertheless, certain members of the sample identified limitations of the instrument. The most common limitation cited was that certain statements were ambiguous or nonspecific to personal circumstances. Phrases such as "less than usual" or "more often" were criticized as being ambiguous, while a statement such as "I am not doing *any* of the (shopping) I would usually do" was regarded as too absolute insofar as it did not admit the possibility of doing "less" (shopping) than usual.

Without exception, respondents elaborated upon their reasons for checking certain SIP items and not others. All sought to describe in more detail their functional health status, suggesting that the SIP is a useful catalyst for obtaining individual-specific profiles of health within an interview setting.

A few respondents commented that the SIP seemed to be applicable to people more disabled than they were. This was

interpreted as verification of the floor effect noted previously in the distribution of SIP scores across categories.

It was also mentioned by respondents that it may be more socially acceptable *not* to check items such as, "I have frequent outbursts of anger at family members, for example, strike at them, scream, throw things at them". In this event SIP data may not reflect the true state of an individual's behavioural dysfunction. Three individuals said the SIP was too lengthy, while another respondent suggested it should contain items related to patient knowledge of specific disease conditions.

I. Evaluation of the SIP by Health Care Professionals

In general, the health care professionals of the Day Hospital familiar with the Sickness Impact Profile (n=8) evaluated the instrument favourably. The SIP was considered to be appropriate for use with the elderly, suitable for inclusion in medical records, easy to comprehend, capable of yielding information from the patient not otherwise readily available, useful as a tool for increasing participation of patients in health care processes, and appropriate as an outcome measure of the Day Hospital program. Table V.17 summarizes these results.

Notably, 62.5% of these health professionals did not consider the SIP to be brief in terms of time required for administration. However, it is questionable whether they

TABLE V.17
EVALUATION OF THE SIP BY HEALTH CARE PROFESSIONALS (N = 8)

CRITERIA (PARAPHRASED)	YES (%)	NO (%)	UNDECIDED (%)
1. Appropriate for use with the elderly.	8(100%)		
2. Suitable for inclusion in medical records.	7(87.5%)	1(12.5%)	
3. Easy to comprehend.	8(100%)		
4. Brief administration time.	3(37.5%)	5(62.5%)	1(12.5%)
5. Yields information from patient not otherwise readily available.	7(87.5%)		1(12.5%)
6. Useful tool for increasing participation of patients in health care process.	7(87.5%)		1(12.5%)
7. Appropriate outcome measure of Day Hospital program.	6(75%)		2(25%)

possessed sufficient information to evaluate the SIP in this regard as they were not directly involved in its administration.

VI. Discussion

In recognition of the need for comprehensive, valid, and reliable measures of perceived health status, expressed in terms of function, and suitable for use with the elderly, a validation study of the Sickness Impact Profile (SIP) was conducted. More specifically, the SIP's descriptive validity, construct validity, test-retest reliability, internal consistency, and utility as an outcome measure of health were assessed. In addition, the relationships between the age, sex, self-rated health, and overall SIP scores of the respondents (N=30) were examined in view of previous studies of perceived health status, and specifically with reference to interpretations of the sex morbidity ratio among the elderly.

As a general statement, it may be said that the SIP was demonstrated to be a reliable and valid instrument, capable of yielding in behavioural terms a comprehensive, detailed description of the total sample, subgroups within it, and individual respondents. Secondly, the relationships between perceived health status and psychological distress were more clearly defined, and confirmed previous research by Blazer & Houpt (1979) that relatively negative perceptions of health are associated with depression, among other factors. Thirdly, the applicability of the SIP to elderly samples, such as the one in this study, was generally supported.

Discussion of the research findings is organized into four sections, each pertaining to an objective of the study.

That is, the descriptive validity of the SIP, the construct validity of the SIP, its utility as an outcome measure of the Day Hospital program, and the interrelationships between the age, sex, objective health status, and self-assessed health of the participants in the study will be discussed as follows.

A. Descriptive Validity of the SIP

The first objective of the study was to assess the descriptive validity of the Sickness Impact Profile with regard to a small group of elderly people attending a geriatric Day Hospital. That is, the extent to which dimension and category scores provided a useful and meaningful qualitative description of the total sample and of different types of patients within it was investigated.

In general, with the exception of the category "Work", the SIP was able to identify and describe in detail the physical and psychosocial functional status of the sample. When participants were divided into groups on the basis of presence or absence of diagnosed depression, the SIP was able to discriminate between these subgroups specifically in terms of category scores in the psychosocial dimension. Concurrently, a descriptive profile of highly prevalent dysfunctions identified by the depressed patients was provided.

As illustrated by the case examples, the SIP appeared to confirm clinical reports of the nature of an individual's

health problems when expressed in terms of function. Given that specific diagnoses and active problems identified by health professionals are not measures of perceived health, and in the case of the problems listed are not always stated with regard to functional status, some discrepancies between SIP profiles and clinical reports were anticipated and did occur. However, ongoing informal feedback from the health professionals of the Day Hospital generally indicated the SIP reinforced existing impressions, or highlighted specific health concerns of patients that may have been de-emphasized in clinical reports.

As previously noted, most members of the sample considered the SIP to give an accurate picture of how they saw their health, and without exception, respondents elaborated upon their reasons for checking certain SIP items and not others. All sought to describe their health in more detail, suggesting the SIP is a useful catalyst for obtaining individual profiles of health within an interview setting. Given the willingness of respondents to share their SIP data with the health professionals of the Day Hospital, it may also be said that the SIP is a useful tool for increasing the participation of patients in health care processes.

In summary, the findings of this research tend to support the claim of Bergner et al. (1981) that "sensitivity of the instrument to different conditions or diagnoses has been tested and results obtained indicate the value of the

SIP in describing similarities of groups of patients and differentiating among these groups" (p. 805). Additionally, interviewer administration of the SIP may facilitate collection of extensive qualititative information supplementary to that directly obtainable from the instrument itself. This information may serve to clarify patient perspectives of health status.

B. Construct Validity of the SIP

The second objective of the study was to assess the construct validity of the SIP in terms of its relationships with indicators of objective health status and patient self-ratings of health.

As a measure of health status, and specifically of perceived health, the construct validity of the SIP was supported by its significant relationships with the number of active problems listed and self-ratings of health. The SIP appeared sensitive not only to the relative severity of dysfunction of the respondents, but also to the nature of health problems experienced. In addition, the SIP seems to represent an extension of the construct of "perceived" or "subjective" health status, insofar as it taps physical as well as psychosocial function. That is, where self-ratings of health have been primarily linked to "feeling-states", the SIP expresses conceptions of health more concretely and comprehensively in terms of performance of activities across a range of areas of living.

C. The SIP as an Outcome Measure of the Day Hospital Program

The third objective of the study was to assess the utility of the SIP as an outcome measure of the Day Hospital program. Given that the SIP was designed "to provide a measure of the effect or outcomes of health care that can be used for evaluation, program planning, and policy formulation" (Bergner et al., 1981, p. 787), its comprehensiveness, demonstrated validity, and overall reliability lend credence to its use for these purposes.

However, it remains uncertain to what extent scores on the SIP vary as a function of health care interventions. As stated by Brook et al. (1977),

> If the influences of natural history of disease, patient characteristics, or other external influences are greater than those of diagnoses and treatment, the measures may not tell much about outcomes, ie., that part of health status directly attributable to changes resulting from the intervention of medical care (p. 9).

Therefore, it may be necessary to assume that observed variations in SIP scores between the first and second occasions of measurement reflect actual changes in respondents' perceived health status which are attributable to the Day Hospital program. Under this assumption, a positive impact of the program is implied by the observation that most respondents had lower overall SIP scores at the time of the second interviews. Secondly, in view of the significant difference observed in the categories Emotional Behaviour, Alertness Behaviour, Body Care and Mobility, and Recreation and Pastimes, it may be stated that improvements

were generally perceived by respondents in the areas of psychosocial function, self-care and mobility, and recreational activities. These improvements, in turn, may reflect emphases within the Day Hospital program. It is also recognized, however, that the SIP is intended to supplement rather than replace other measures of health status, and as such it does not claim to reflect all areas in which the impact of an program may be evident.

In summary, in view of its aforementioned advantages, as well as its general acceptability to patients and professional staff, it is suggested that the SIP merits serious consideration as an outcome measure of health, specifically from the viewpoint of the patient, and supplementary to more traditional outcome measures.

D. Age and Sex as Predictors of Perceived Health

The fourth objective of the study was to examine the relationships between the age, sex, objective health status, and self-assessed health of the participants in the study.

The lack of significant relationships between the age and sex of the respondents and measures of objective and subjective health status has been noted, and the hypothesis advanced that this may be due to the homogeniety of the sample with regard to health status. It is also possible that when perceived health is operationalized in terms of specific areas of performance, the influence of any social norms becomes secondary to perceptions of actual performance

of activities. While it is not feasible to more than speculate within the confines of the present study, further investigation of the relationships between perceived health, as expressed within a performance conception of health, and the sex morbidity ratio appears warranted.

Notably, however, the finding of Blazer & Houpt (1979) that depression, among other factors, is a more important predictor of self-rated health than the sex or age of respondents was replicated in this study. This is also consistent with Tessler & Mechanic's (1978) observation that "psychological distress and persons' perceptions of their physical well-being" are significantly correlated. Indeed, this relationship was found to be more specific insofar as depression was related to relatively negative self-ratings of health and psychosocial dysfunction as measured by the SIP.

E. Concluding Remarks

In conclusion, it is apparent that utilization of behaviourally based measures of perceived health status such as the Sickness Impact Profile represents a significant advancement in measurement of the self-assessed health of the elderly. Communication of perceptions of health status in behavioural terms facilitates development of intervention plans consistent with individual patient priorities, and in so doing indicates respect for the patient perspective. Further, gaining insight into patient values may serve to

clarify perceived needs for services as well as encourage further steps toward understanding the interrelationships between health, behaviour, and aging.

REFERENCES

Allen, M. J. & Yen, W. M. *Introduction to measurement theory*. Monterey, California: Brooks/Cole Publishing Company, 1979.

Atchley, R. C. Selected social and psychological differences between men and women in later life. *Journal of Gerontology*, 1976, **31**, 204-211.

Baumann, B. Diversities in conceptions of health and physical fitness. *Journal of Health and Social Behaviour*, 1961, **2**, 39-46.

Bay, K. S., Leatt, P., & Stinson, S. M. A patient classification system for long-term care. *Medical Care*, 1982, **20**, 468-488.

Bay, K. S., Overton, P., Harrison, F. P., Stinson, S. M., & Hazlett, S. L. *Patient classification by types of care*. A research report on the development and validation of the PCTC system. The University of Alberta: March, 1979.

Beeson, D. Women in studies of aging: a critique and suggestion. *Social Problems*, 1975, **23**, 52-59.

Berg, R. L., Hallauer, D. S., & Berk, S.N. Neglected aspects of the quality of life. *Health Services Research*, 1976, 11, 391-395.

Bergner, M., Bobbitt, R. A., Carter, W. B., & Gilson, B. S. The Sickness Impact Profile: development and final revision of a health status measure. *Medical Care*, 1981, 19, 787-805.

Bergner, M. & Gilson, B. S. The Sickness Impact Profile: the relevance of social science to medicine. In L. Eisenberg & A. Kleinman (Eds.), *The relevance of social science for medicine*. Dordrecht, Holland: D. Riedel Publishing Company, 1981, 135-150.

Bergner, M., Bobbitt, R. A., Kressel, S., Pollard, W. E., Gilson, B. S., & Morris, J. R. The Sickness Impact Profile: conceptual formulation and methodology for the development of a health status measure. *International Journal of Health Services*, 1976a, 6, 393-415.

Bergner, M., Bobbitt, R. A., Pollard, W. E., Martin, D. P., & Gilson, B. S. The Sickness Impact Profile: validation of a health status measure. *Medical Care*, 1976b, 14, 57-67.

Birren, J. E., Cunningham, W. R., & Yamamoto, K. The

psychology of adult development and aging. *Annual Review of Psychology*, 1983, **34**, 543-575.

Blazer, D. G. & Houpt, J. F. Perception of poor health in the healthy older adult. *Journal of the American Geriatrics Society*, 1979, **27**, 330-334.

Branch, L. G., Jette, A., Evaswick, C., Polansky, M., Rowe, G., & Diehr, P. Toward understanding elders' health service utilization. *Journal of Community Health*, 1981, **7**, 80-92.

Branch, L. G. Functional abilities of the elderly: an update on the Massachusetts Health Care Panel Study. In Haynes, S. G. & Feinleib, M. (Eds.), *Second conference on the epidemiology of aging*. Washington, D. C.: U. S. Department of Health and Human Services, NIH Pub. No. 80-969, 1980.

Branch, L. G. & Jette, A. M. The Framingham Disability Study: I. Social disability among the aging. *American Journal of Public Health*, 1981, **71**, 1202-1210.

Brook, R. H., Davis-Avery, A., Greenfield, S., et al. Assessing the quality of medical care using outcome measures: an overview of the method. *Medical Care*, 1977, **15**, Suppl. 9, 1-84.

Campbell, D. T. & Fiske, D. W. Convergent and discriminant validation by the multitrait-multimethod matrix. In Jackson, D. N. & Messick, S. *Problems in human assessment*. New York: McGraw-Hill Book Company, 1967, 124-132.

Carter, W. B. & Deyo, R. A. The impact of questionnaire research on clinical populations: a dilemma in review of human subjects research resolved by a study of a study. *Clinical Research*, 1981, **29**, 287-295.

Carter, W. B., Bobbitt, R. A., Bergner, M., & Gilson, B. S. Validation of an interval scaling: the Sickness Impact Profile. *Health Services Research*, 1976, 11, 516-528.

Chambers, L. W. Health program review in Canada: measurement of health status. *Canadian Journal of Public Health*, 1982, **73**, 26-34.

Chappell, N. L. Measuring functional ability and chronic health conditions among the elderly: a research note on the adequacy of three instruments. *Journal of Health and Social Behaviour*, 1981, **22**, 90-102.

Chappell, N. L. & Havens, B. Old & female: testing the double jeopardy hypothesis. *Sociological Quarterly*,

1980, **21**, 157-171.

Chen, M. K. & Bryant, B. E. The measurement of health: a critical and selective overview. *International Journal of Epidemiology*, 1975, **4**, 257-264.

Cleary, P. D., Mechanic, D., & Greenley, J. R. Sex differences in medical care utilization: an empirical investigation. *Journal of Health and Social Behaviour*, 1982, **23**, 106-119.

Conn, J., Bobbitt, R. A., & Bergner, M. *Administration procedures and interviewer training for the Sickness Impact Profile*. Seattle, Washington: Department of Health Services, University of Washington, 1978.

Cook, T. D. & Campbell, D. T. *Quasi-experimentation: design and analysis issues for field settings*. Chicago: Rand McNally College Publishing Company, 1979.

Cronbach, L. J. Test validation. In Thorndike, R. L. (Ed.), *Educational Measurement*. Washington, D. C.: American Council on Education, 1971, 443-507.

Cronbach, L. J. & Gleser, G. C. Assessing similarity between profiles. *Psychological Bulletin*, 1953, **50**, 456.

Cronbach, L. J. Coefficient alpha and the internal structure of tests. *Psychometrika*, 1951, **16**, 297-334.

Davis, M. A. Sex differences in reporting osteoarthritic symptoms: a sociomedical approach. *Journal of Health and Social Behaviour*, 1981, **22**, 298-310.

Deyo, R. A., Inui, T. S., Leininger, J. D., & Overman, S. S. Measuring functional outcomes in chronic disease: a comparison of traditional scales and a self-administered health status questionnaire in patients with rheumatoid arthritis. *Medical Care*, 1983, **21**, 180-192.

Deyo, R. A., Inui, T. S., Leininger, J., & Overman, S. S. Physical and psychosocial function in rheumatoid arthritis: clinical use of a self-administered health status instrument. *Archives of Internal Medicine*, 1982, **142**, 879-882.

Dohrenwend, B. & Dohrenwend, B. Reply to Gove and Tudor's comment on 'sex differences in psychiatric disorders'. *American Journal of Sociology*, 1977, **82**, 1336-1345.

Dulude, L. *Women and aging: a report on the rest of our lives*. Ottawa: Canadian Advisory Council on the Status of Women, 1978.

Edwards, J. & Klemmack, D. Correlates of life satisfaction. *Journal of Gerontology*, 1973, **28**, 497-502.

Elias, M. F., Elias, P. K., & Elias, J. W. Health and behaviour. In, *Basic processes in adult developmental psychology*. Saint Louis: The C. V. Mosby Company, 1977, 221-255.

Fabrega, H. The need for an ethnomedical science. *Science*, 1975, **189**, 969-975.

Ferraro, K. F. Self-ratings of health among the old and the old-old. *Journal of Health and Social Behaviour*, 1980, **21**, 377-383.

Fillenbaum, G. G. Social context and self-assessments among the elderly. *Journal of Health and Social Behaviour*, 1979, **20**, 45-51.

Friedsam, H. J. & Martin, H. W. A comparison of self and physician's health ratings in an older population. *Journal of Health & Social Behaviour*, 1964, **4**, 179-183.

George, L. K. & Bearon, L. B. *Quality of life in older persons: meaning and measurement*. New York: Human Sciences Press, 1980.

Giesen, C. B. & Datan, N. The competent older woman. In N. Datan & N. Lohmann (Eds.), *Transitions of aging*. New York: Academic Press, 1980, 57-72.

Gilson, B. S., Bergner, M., Bobbitt, R. A., Kressel, S., Pollard, W. E., & Vesselago, M. The Sickness Impact Profile: development of an outcome measure of health care. *American Journal of Public Health*, 1975, **65**, 1304-1310.

Government of Canada. *Canadian Governmental Report on Aging*. Ottawa: Minister of Supply & Services Canada, June 1982.

Government of Canada. *The health of Canadians: report of the Canada Health Survey*. Ottawa: Minister of Supply & Services Canada, 1981.

Graney, M. J. & Zimmerman, R. M. Causes and consequences of health self-report variations among older people. *International Journal of Aging and Human Development*, 1980-81, **12**, 291-300.

Gutman, G. The elderly at home and in retirement housing: a comparative study of health problems, functional difficulties, and support service needs. In V. W. Marshall (Ed.), *Aging in Canada: social perspectives*. Don Mills: Fitzhenery & Whiteside, 1980, 189-200.

Haber, L. D. Disabling effects of chronic disease and impairment-II: functional capacity limitations. *Journal of Chronic Diseases*, 1973, **26**, 127-151.

Havens, B. J. Differentiation of unmet needs using analysis by age/sex cohorts. In Marshall, V. W. (Ed.), *Aging in Canada: social perspectives*. Don Mills, Ontario: Fitzhenry & Whiteside, 1980, 215-221.

Heyman, D. K. & Jeffers, F. C. Effect of time lapse on consistency of self-health and medical evaluations of elderly persons. *Journal of Gerontology*, 1964, **18**, 160-164.

Hunt, S. M., McKenna, S. P., McEwen, J., Backett, E. M., Williams, J., & Papp, E. A quantitative approach to perceived health status: a validation study. *Journal of Epidemiology and Community Health*, 1980, **34**, 281-286.

Jette, A. M. & Branch, L. G. The Framingham Disability Study: II. Physical disability among the elderly. *American Journal of Public Health*, 1981, **71**, 1211-1216.

Jette, A. M. Health status indicators: their utility in chronic-disease evaluation research. *Journal of Chronic Diseases*, 1980, **33**, 567-579.

Jette, A. M. & Deniston, O. L. Inter-observer reliability of a functional status assessment instrument. *Journal of Chronic Diseases*, 1978, **31**, 573-580.

Johnson, M. Self-perception of need among the elderly: an analysis of illness behaviour. *The Sociologic Review*, 1972, **20**, 521-531.

Kaplan, R. M., Bush, J. W., & Berry, C. C. Health status: types of validity and the index of well-being. *Health Services Research*, 1976, **11**, 478-507.

Katz, S., Ford, A. B., Moskowitz, R. W., Jackson, B. A., & Jaffe, M. W. Studies of illness in the aged: the index of ADL: a standardized measure of biological and psychosocial function. *American Medical Association Journal*, 1963, **185**, 914-919.

Katz, S., Downs, T. D., Cash, H. R., & Grotz, R. C. Progress in the development of the index of ADL. *The Gerontologist*, 1970, **Spring**, 20-30.

Kaufert, J. M., Green, S., Dunt, D. R., Corkhill, R., Creese, A. L., & Locker, D. Assessing functional status among elderly patients: a comparison of questionnaire and service provider ratings. *Medical Care*, 1979, **17**,

807-817.

Kline, C. The socialization process of women: implications for a theory of successful aging. *The Gerontologist*, 1975, **15**, 486-492.

Larson, R. Thirty years of research on the subjective well-being of older Americans. *Journal of Gerontology*, 1978, **33**, 109-125.

Larue, A., Bank, L., Jarvik, L., & Hetland, M. Health in old age: how do physician's and self-ratings compare? *Journal of Gerontology*, 1979, **34**, 687-691.

Lawton, M. P. The functional assessment of elderly people. *Journal of the American Geriatrics Society*, 1971, **19**, 465-481.

Linn, M. W. Studies in rating the physical, mental, and social dysfunction of the chronically ill aged. *Medical Care*, 1976, **14**, 119-125.

Maddox, G. L. & Douglass, E. B. Self-assessment of health: a longitudinal study of elderly subjects. *Journal of Health and Social Behaviour*, 1973, **14**, 87-93.

Maddox, G. L. Self-assessment of health status: a

longitudinal study of selected elderly subjects. *Journal of Chronic Diseases*, 1964, **17**, 449-460.

Maddox, G. L. Some correlates of differences in self-assessment of health status among the elderly. *Journal of Gerontology*, 1962, **17**, 180-185.

Mancini, J. A. & Quinn, W. H. Dimensions of health and their importance for morale in old age: a multivariate examination. *Journal of Community Health*, 1981, **7**, 118-128.

Martin, D. P., Gilson, B. S., Bergner, M., Bobbitt, R. A., Pollard, W. E., Conn, J. R., & Cole, W. A. The Sickness Impact Profile: potential use of a health status instrument for physician training. *Journal of Medical Education*, 1976, **51**, 942-944.

Martini, C. J. & McDowell, I. Health status: patient and physician judgments. *Health Services Research*, 1976, **11**, 508-515.

Mausner, J. S. & Bahn, A. K. *Epidemiology: an introductory text*. Philadelphia: W. B. Saunders Company, 1974.

Marascuilo, L. A. & McSweeney, M. *Non-parametric and distribution free methods for the social sciences*.

Monterey, California: Brooks/Cole Publishing Company, 1977.

McCrae, R. R., Bartone, P. T., & Costa, P. T. Age, anxiety, and self-reported health. *International Journal of Aging and Human Development*, 1976, **7**, 49-58.

McIntyre, N. The principles of the problem oriented medical record. In J. C. Petrie & N. McIntyre (Eds.), *The problem oriented medical record*. London: Churchill Livingstone, 1979, 35-71.

Mechanic, D. Sex, illness behaviour, and the use of health services. *Social Science & Medicine*, 1978, **12**, 207-214.

Moore, J. T. & Fillenbaum, G. G. Change in functional disability of geriatric patients in a family medicine program: implications for patient care. *Journal of Family Practice*, 1981, **12**, 59-66.

Mossey, J. M. & Shapiro, E. Self-rated health: a predictor of mortality among the elderly. *American Journal of Public Health*, 1982, **72**, 800-808.

Muller, C. Health status and survival needs of the elderly. *American Journal of Public Health*, 1982, **72**, 789-790.

Nathanson, C. Sex, illness, and medical care: a review of data, theory, and method. *Social Science and Medicine*, 1977, **11**, 13-25.

Nathanson, C. Illness and the feminine role: a theoretical review. *Social Science & Medicine*, 1975, **9**, 57-62.

Nie, N. H., Hull, C. H., Jenkins, J. G., Steinbrenner, K., & Bent, D. H. *Statistical package for the social sciences*. New York: McGraw-Hill, 1975.

Palmore, E. & Luikart, C. Health and social factors related to life satisfaction. *Journal of Health and Social Behaviour*, 1972, **13**, 68-80.

Parkerson, G. R., Gehlbach, S. H., Wagner, E. H., James, S. A., Clapp, N. E., & Muhlbaier, L. H. The Duke-UNC health profile: an adult health status instrument for primary care. *Medical Care*, 1981, **19**, 806-823.

Patrick, D. L., Bush, J. W., & Chen, M. M. Toward an operational definition of health. *Journal of Health & Social Behaviour*, 1973, **14**, 6-23.

Payne, B. & Whittington, F. Older women: an examination of popular stereotypes and research evidence. *Social Problems*, 1976, **23**, 488-504.

Phillips, D. & Segal, B. Sexual status and psychiatric symptoms. *American Sociological Review*, 1969, **34**, 58-72.

Pollard, W. E., Bobbitt, R. A., & Bergner, M. Examination of variable errors of measurement in a survey-based social indicator. *Social Indicators Research*, 1978, 279-301.

Pollard, W. E., Bobbitt, R. A., Bergner, M., Martin, D. P., & Gilson, B. S. The Sickness Impact Profile: reliability of a health status measure. *Medical Care*, 1976, **14**, 146-155.

Reynolds, W. J., Rushing, W. A., & Miles, D. L. The validation of a function status index. *Journal of Health and Social Behaviour*, 1974, **15**, 271-283.

Schwenger, C. W. & Sayers,. L. A. A Canadian survey by public health nurses of the health and living conditions of the aged. *American Journal of Public Health*, 1971, **61**, 1189-1195.

Shanas, E. & Maddox, G. L. Aging, health, and the organization of health resources. In Binstock, R. H. & Shanas, E. *Handbook of aging and the social sciences*. New York: Van Nostrand Reinhold Company, 1976.

Shanas, E. The psychology of health. In B. L. Neugarten (Ed.), *Middle age and aging: a reader in social psychology*. Chicago: University of Chicago Press, 1968, 212-219.

Siegler, I. C., Nowlin, J. B., & Blumenthal, J. A. Health and behaviour: methodological considerations for adult development and aging. In, Poon, L. W. (Ed.), *Aging in the 1980's: psychological issues*. Washington, D. C.: American Psychological Association, 1980.

Snider, E. Explaining life satisfaction among the elderly. *Social Science Quarterly*, 1980, 61, 253-263.

Social Sciences and Humanities Research Council of Canada. *Ethical guidelines for the institutional review committees for research with human subjects*. Ottawa: SSHRC of Canada, September 1981.

Spreitzer, E. & Snyder, E. Correlates of life satisfacton among the aged. *Journal of Gerontology*, 1974, 29, 454-458.

Statistics Canada, Census Characteristics Division. *Canada's Elderly*. Ottawa: Statistics Canada, 1979.

Statistics Canada. *Census of Canada 1976*. Volume 2, part 1.

Population: demographic characteristics. Ottawa: Statistics Canada, 1976.

Stenback, A., Kumpulainen, M., & Vauhkonen, M-L. Illness and health behaviour in septuagenarians. *Journal of Gerontology*, 1978, **33**, 57-61.

Stewart, A., Ware, J. E., & Brook, R. H. The meaning of health: understanding functional limitations. *Medical Care*, 1977, **15**, 939-952.

Stone, L. O. & Fletcher, S. *A profile of Canada's older population*. Montreal: The Institute for Research on Public Policy, 1980.

Suchman, E. A., Phillips, B. S., & Streib, B. F. An analysis of the validity of health questionnaires. *Social Forces*, 1958, **36**, 223-232.

Sullivan, D. F. Conceptual problems in developing an index of health. *Health and Vital Statistics*, Series 2, No. 17. National Centre for Health Statistics. Washington, D. C.: U. S. Government Printing Office, 1966.

Tissue, T. Another look at self-rated health among the elderly. *Journal of Gerontology* , 1972, **27**, 91-94.

Tessler, R. & Mechanic, D. Psychological distress and perceived health status. *Journal of Health and Social Behaviour*, 1978, 19, 254-262.

Verbrugge, L. M. Sex differentials in morbidity and mortality in the United States. *Social Biology*, 1976a, 23, 275-296.

Verbrugge, L. M. Females and illness: recent trends in sex differences in the United States. *Journal of Health and Social Behaviour*, 1976b, 17, 387-403.

Wan, T. T. H. Predicting self-assessed health status: a multivariate approach. *Health Services Research*, 1976, 11, 464-477.

Wilkins, R. *Health status in Canada: 1926-1976*. Montreal: Institute for Research on Public Policy, 1980.

APPENDIX A

Letter of Introduction to Sample

APPENDIX A: Letter of Introduction to Sample

Dear :

This letter is an invitation to participate in a research study to be conducted at the Day Hospital.

The purpose of the study is to see what useful information can be gained from using a "self-assessment of health". Often in the past we have depended on medical records to tell us how healthy people are, but more recently health care professionals are finding that people can tell us a great deal about their own state of health. The Sickness Impact Profile (SIP) is a questionnaire that is designed to collect information from people concerning their own health. It is believed that the information would provide an important supplement to medical records. It is also believed that the SIP may be especially useful in providing information on how older people see their health, and I have proposed to do a study to demonstrate the usefulness of the SIP for this purpose. To conduct the study, I need the cooperation of a group of elderly people and so I am inviting you to become a member of the study sample.

If you agree to participate, I will arrange a time for us to meet at the Day Hospital when it is convenient for you. I will ask you to respond to the items on the SIP

questionnaire on two occasions. About three weeks will pass between my first and second interviews with you. Both times you shall be asked to state whether or not each item on the questionnaire describes you and is related to your health.

Examples of items on the SIP are:

"I walk up or down stairs only with assistance from someone else".

"I make difficult moves with help, for example, getting into or out of cars, bathtubs".

"I often express concern over what might be happening to my health".

There are 136 items in total on the questionnaire. Each interview will require about 30 minutes of your time.

Immediately following the second interview, I will ask you to give your opinions about the SIP questionnaire so that I may know how useful you considered it to be in describing how you see your health.

Also, in order to determine the relationship between your responses to the SIP and the assessment of your health made by the professional staff of the Day Hospital, I request your permission to review your medical record there.

I assure you that all personal information given, as well as all information obtained from medical records, will be regarded as confidential. The information will be coded

numerically so as to ensure that anonymity is maintained. Following complete analysis of the data, all information pertaining to you as an individual will be destroyed. If you have any questions about the procedures to be used, I will be pleased to answer them.

In closing, I thank you for considering my invitation to participate in the study. A consent form is provided below in which you may indicate if you are willing or not to participate in it.

Sincerely,

Allison L. McKinnon

Are you willing to participate as member of the sample in the research study described in the above letter?

YES_____ NO_____

SIGNATURE: _____

DATE: _____

WITNESS: _____

APPENDIX B

Letter of Consent for Use of the Sickness Impact Profile

UNIVERSITY OF WASHINGTON
SEATTLE, WASHINGTON 98195
SCHOOL OF PUBLIC HEALTH AND COMMUNITY MEDICINE
Department of Health Services SC-37
Community Medicine (206) 543-7952

May 26, 1982

Ms. Allison L. McKinnon
10805 - 73 Avenue
Edmonton, Alberta
CANADA
T6E 1C8

Dear Ms. McKinnon:

The research that you propose in your letter of May 5 seems completely appropriate for the Sickness Impact Profile. Please consider this letter as an indication of permission to use the SIP in your Master's thesis research. I would, however, appreciate receiving a detailed description of the proposed research and, when it is completed, a copy of the thesis itself.

Sincerely yours,

Marilyn Bergner, Ph.D.
Associate Professor

MB:lah
Enclosures

APPENDIX C

Participant Evaluation of the SIP

APPENDIX C: Participant Evaluation of the SIP

Now that you have had an opportunity to answer the questions in the questionnaire known as the Sickness Impact Profile, do you think it gives an accurate picture of how you see your health at this time?

YES NO

Comments: _____

_____

_____

Did you have difficulty in understanding the instructions for responding to the SIP?

YES NO

Comments: _____

_____

_____

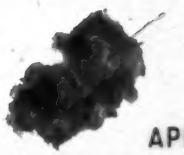
Do you think the information you provided in answering the Sickness Impact Profile should be shared with the health professionals directly concerned with your program at the Day Hospital?

YES NO

Comments: _____

_____

_____



APPENDIX D

Health Care Professionals Questionnaire

APPENDIX D: Health Care Professionals Questionnaire

Evaluation of the Sickness Impact Profile

Now that you are familiar with the Sickness Impact Profile, do you consider it to be:

	YES	NO	UNDECIDED
1. Appropriate for use with an elderly population?			
2. Suitable for inclusion in medical or health records?			
3. Easy to comprehend?			
4. Brief with regard to administration time required?			
5. Capable of yielding information about the patient's perspective of his health not otherwise readily available from existing assessment procedures?			
6. A useful tool for increasing participation of patients in health care processes?			
7. Appropriate for use as an outcome measure of the Day Hospital program?			

APPENDIX E

Highly Prevalent Behavioural Dysfunctions of Sample, as Measured by the SIP

APPENDIX E
HIGHLY PREVALENT BEHAVIOURAL DYSFUNCTIONS OF SAMPLE, AS MEASURED BY THE SIP

CATEGORY	SIP ITEM (PARAPHRASED)
SR	...sit during much of the day.
	...sleep less at night.
HM	...not doing heavy work around house.
M	...stay home most of the time.
SI	...going out less to visit people.
	...doing fewer social activities with groups.
A	...walk shorter distances...rest often.
	...do not walk up or down hills.
	...walk more slowly.
RP	...going out for entertainment less.
	...doing fewer community activities.
E	...eating special or different food.

APPENDIX F

Behavioural Dysfunctions Not Identified in Sample, as Measured by the SIP

APPENDIX F: Behavioural Dysfunctions Not Identified in Sample, as Measured by the SIP

CATEGORY	SIP ITEM (PARAPHRASED)
BCM	...use bedpan with assistance...
M	...staying in bed most of the time...
SI	...frequent outbursts of anger at family members... ...isolate myself from the rest of the family... ...refuse contact with family members...
A	...walk only with help from someone...
C	...speech understood only by a few people... ...have difficulty speaking...
W	...not working at all... ...doing part of my job at home... ...not accomplishing as much as usual... ...act irritable toward work associates... ...doing only light work... ...work only for short periods of time... ...do not do my job as carefully as usual...
E	...eat no food at all...taking fluids... ...pick or nibble at my food... ...feed myself with help from someone else... ...do not feed myself at all... ...eating no food...tubes or intravenous fluids...

APPENDIX G

Highly Prevalent Behavioural Dysfunctions of Depressed Patients, as Measured by the SIP.

APPENDIX G

HIGHLY PREVALENT BEHAVIOURAL DYSFUNCTIONS OF DEPRESSED PATIENTS, AS MEASURED BY THE SIP (n=15)

CATEGORY	SIP ITEM (PARAPHRASED)	NO. (%) CHECKING ITEM
SR	...sit during much of the day.	8 (53.3%)
	...sleep less at night.	8 (53.3%)
EB	...act nervous or restless.	12 (80.0%)
	...talk about the future in a hopeless way	8 (53.3%)
BCM	...stand only for short periods.	8 (53.3%)
	...kneel, stoop...by holding on to something.	10 (66.7%)
	...change position frequently.	8 (53.3%)
HM	...work around house only for short periods.	9 (60.0%)
	...not doing heavy work around house.	10 (66.7%)
M	...stay home most of the time.	9 (60.0%)
	...stay away from home only for brief periods.	10 (66.7%)
SI	...going out less to visit people.	12 (80.0%)
	...doing fewer social activities with groups.	10 (66.7%)
	...often express concern over...my health.	9 (60.0%)
A	...walk shorter distances...rest often.	10 (66.7%)
	...do not walk up or down hills.	11 (73.3%)
	...go up and down stairs more slowly.	11 (73.3%)
	...walk more slowly.	12 (80.0%)
AB	...difficulty reasoning and solving problems.	9 (60.0%)
	...make more mistakes than usual.	8 (53.3%)
	...difficulty...with concentration and thinking.	9 (60.0%)
C	...trouble writing or typing.	8 (53.3%)
RP	...going out for entertainment less.	12 (80.0%)
	...cutting down on some usual inactive pastimes.	8 (53.3%)
	...doing fewer community activities.	10 (66.7%)
E	...eating much less than usual.	8 (53.3%)
	...eating special or different food.	8 (53.3%)

APPENDIX H

Behavioural Dysfunctions of All Patients with "Poor" Self-Rated Health, as Measured by the SIP

APPENDIX H: Behavioural Dysfunctions of All Patients with 'Poor' Self-Rated Health, as Measured by the SIP

CATEGORY	SIP ITEM (PARAPHRASED)
EB	...say how bad or useless I am... ...keep rubbing or holding areas of my body that hurt...
BCM	...stand only for short periods...
HM	...not doing heavy work around house...
SI	...going out less to visit people.
RP	...going out for entertainment less...